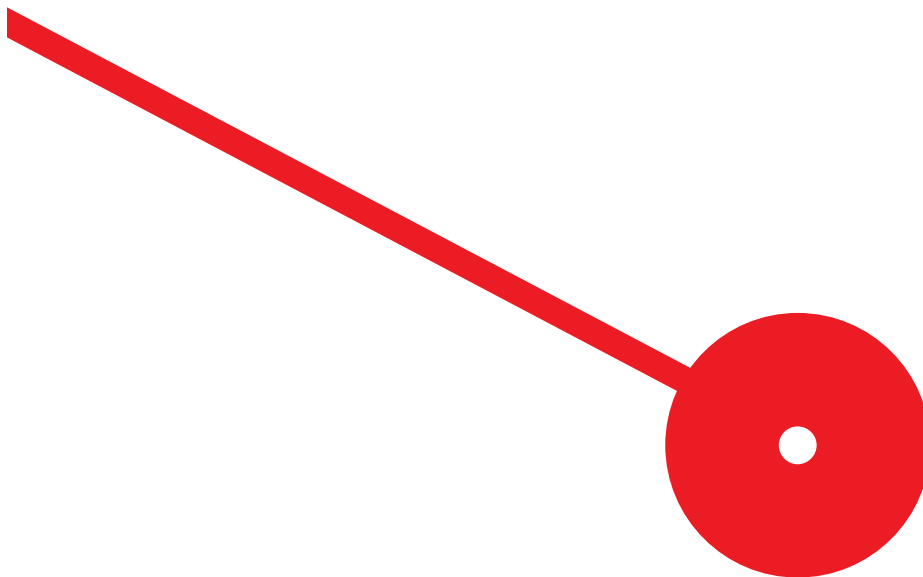




Big Data no comércio eletrónico: utilização de *cookies* e questões de privacidade na recolha de dados pessoais

Joana Luís da Silva

07/2020



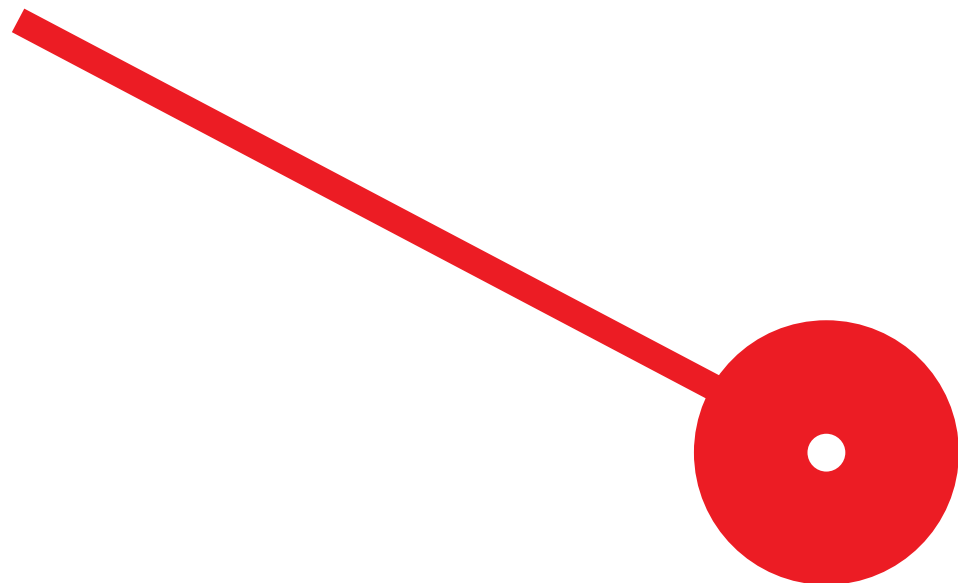


Big Data no comércio eletrónico: utilização de *cookies* e questões de privacidade na recolha de dados pessoais

Joana Luís da Silva

Dissertação de Mestrado apresentado ao Instituto Superior de Contabilidade e Administração do Porto para a obtenção do grau de Mestre em Negócio Eletrónico, sob orientação da Doutora Ana Isabel Rojão Lourenço Azevedo.

Joana Luís da Silva. Big Data no comércio eletrónico:
utilização de *cookies* e questões de privacidade na
recolha de dados pessoais
07/2020



Agradecimentos

A concretização desta dissertação só foi possível com o apoio de várias pessoas que de certa forma acompanharam o meu percurso académico. Foram diversos os contributos e expresso aqui os meus sinceros agradecimentos.

Em primeiro lugar, quero agradecer de forma muito especial à minha orientadora, Professora Doutora Ana Isabel Rojão Lourenço Azevedo, pela orientação dada ao longo desta etapa e por toda a sua disponibilidade, atenção, profissionalismo e apoio prestado na elaboração desta dissertação.

Ao ISCAP, pela formação que me proporcionou e a todos os docentes que me acompanharam ao longo deste mestrado e que contribuíram para atingir esta etapa.

À minha Mãe, Pai, Irmãos e restantes familiares, pelo apoio incondicional, incentivo, paciência e força dada para ultrapassar obstáculos que foram surgindo.

Aos meus grandes companheiros da licenciatura e de mestrado, Xana, Raquel, Rui e Tiago por todos os bons momentos, diversão e amizade que, decerto, nos marcaram a todos.

Ao meu padrinho, *Rot*, pela paciência, incentivo e calma em todas as horas.

Ao meu Manel, que sempre que me fui abaixo, me lembrou a mulher cheia de garra que sou e me encheu o ego sem medos nem vergonhas nos seus longos discursos motivacionais.

Por fim, mas não menos importante, aos restantes amigos e colegas que o mestrado me deu e pelos momentos partilhados que recordarei para sempre.

Resumo:

O crescimento das redes sociais, a proliferação do comércio eletrônico e a constante invasão de dispositivos móveis fez com que a quantidade de dados aumentasse drasticamente. Big Data refere-se a grandes conjuntos de dados caracterizados por grandes volumes e grande variedade, que têm origem em várias fontes, sendo gerados a alta velocidade. Este fenómeno fez com que o mundo dos negócios se alterasse e estimulou as empresas de comércio eletrônico a utilizarem este tipo de dados, analisando, por exemplo, os dados dos perfis dos seus consumidores e o respetivo perfil de navegação para definir, em tempo real, produtos a serem oferecidos aos seus clientes. O objetivo principal desta dissertação foi o de analisar em que medida os Big Data estão presentes e são utilizados no contexto de comércio eletrônico, quer do ponto de vista dos utilizadores, quer do ponto de vista das organizações. Pretende-se também esclarecer algumas questões éticas e legais no contexto dos Big Data. Esta dissertação caracteriza-se como um *Estudo de Caso* com tipologia *múltipla* do tipo *descritivo* e baseou-se em duas grandes fontes de recolha de dados. A primeira foi um questionário *online* que teve como objetivo avaliar o nível de consciencialização do utilizador em relação à partilha e recolha dos seus dados em ambiente de comércio eletrônico. Partiu-se de uma amostra *não-probabilística* de tipo bola de neve. As respostas válidas dessa amostra (160) foram analisadas numa abordagem *quantitativo-qualitativa*. A segunda foi a ferramenta de pesquisa da Cookiepedia, a qual permitiu uma análise dos *cookies* de um conjunto de *sites* de comércio eletrônico, obtidos através de uma amostra de *conveniência*. Foram analisados a Política de Privacidade de cada *site*, os tipos e objetivos dos *cookies* presentes nesses *sites*, rematando com uma breve lista discriminada dos mesmos. Com este estudo foi possível concluir que os utilizadores *online*, apesar de não concordarem com as condições dadas pelos *sites* de comércio eletrônico para a sua utilização, que os continuam a utilizar e que os utilizadores *online* não têm conhecimento de alguns dados considerados como pessoais. Este estudo conclui também que a maioria dos *cookies* utilizados em *sites* de comércio eletrônico se destinam a publicidade *online*, que é a utilização mais comum dos chamados *Third Party Cookies*.

Palavras chave: RGPD, *Cookies*, Estudo de Caso, Questionário, Cookiepedia

Abstract:

The growth of social media, the proliferation of e-commerce and the constant expansion of mobile devices worldwide has led to an exponential increase of Big Data. Big Data refers to huge data sets characterized by large volumes and large variety, as they originate from various data sources, and are generated at high speed. This phenomenon has changed the business world and forced e-commerce companies to use their consumer's profile data and navigation profile to define, in real time, products to be offered to their customers. The main objective of this dissertation is to analyse the extent to which Big Data is present and used in the context of electronic commerce, both from user's and organisation's point of view, whilst also clarifying some ethical and legal issues. This dissertation is characterized as a descriptive case study and is based on two major sources of data collection. Firstly, with an online questionnaire that seeks to assess the user's level of awareness of sharing and collecting their data in an e-commerce environment, which has generated a non-probabilistic sample. The valid responses of this sample (160) were analysed in a quantitative-qualitative approach using descriptive statistics. The second source is Cookiepedia's research tool, for a sample of convenience, where the Privacy Policy of each, the types and objectives of cookies present in those websites is analysed, ending with a brief list of them. With this study it was possible to conclude that online users, although they do not agree with the conditions given by e-commerce sites for their use, continue to use them and that online users are not aware of some data considered personal. This study also concludes that the majority of cookies used on e-commerce sites are for online advertising, which is the most common use of so-called Third Party Cookies.

Key words: GDPR, *Cookies*, Case Study, Survey, Cookiepedia

Índice geral

Capítulo I - Introdução	1
Capítulo II – Referencial teórico	5
2.1 Conceito de Big Data	6
2.2 Os 6 V's dos Big Data	7
2.3 Tecnologias no âmbito dos Big Data	11
2.3.1 Apache Hadoop	11
2.3.2 MapReduce	12
2.3.3 NoSQL	12
2.4 Benefícios e desafios, oportunidades e riscos dos Big Data	12
2.5 Questões éticas e legais no contexto dos Big Data	15
2.5.1 Proteção de dados nos Big Data e no comércio eletrónico	16
2.5.1.1 Informação recolhida pelos <i>sites</i>	17
2.5.1.2 Principais formas pelas quais as empresas <i>online</i> recolhem informação sobre os utilizadores	18
2.5.2 Regulamento Geral de Proteção de Dados Pessoais na Europa	20
2.5.2.1 Direito à Portabilidade dos Dados	20
2.5.2.2 Direito a Ser Esquecido	21
2.6 Rastreamento <i>Online</i>	22
2.6.1 <i>Cookies</i>	23
2.6.1.1 Tipos de <i>Cookies</i>	24
2.6.1.2 Objetivos dos <i>cookies</i>	24
Capítulo III – Metodologia	27
3.1 O Estudo de Caso	28
3.1.1 Caso Presente	30
3.2 Abordagem à investigação e métodos de recolha de dados	31
Capítulo IV – Resultados	40

4.1.	Análise do questionário <i>online</i>	41
4.1.1.	Caracterização da Amostra	41
4.1.2.	Nível de conhecimento do utilizador online em relação ao conceito de dados pessoais	43
4.1.3.	Avaliação do comportamento do utilizador perante a partilha dos seus dados	45
4.1.4.	Nível de conhecimento perante os seus direitos inerentes no RGDP ...	46
4.1.5.	Nível de concordância perante as opções que os <i>sites</i> oferecem para a sua utilização.....	48
4.1.6.	Comentários adicionais.....	50
4.2.	Análise dos <i>cookies</i> em <i>sites</i> de comércio eletrónico	51
4.2.1.	Políticas de privacidade dos <i>sites</i> de comércio eletrónico	51
4.2.2.	Tipo de <i>cookies</i> presentes nos <i>sites</i> de comércio eletrónico	64
4.2.3.	<i>Cookies</i> presentes nos <i>sites</i> de comércio eletrónico	66
	Capítulo V –Discussão dos resultados	74
	Capítulo VI – Conclusão	78
	Referências	82
	Apêndices.....	87
	Apêndice I – Questionário	88
	Anexos.....	92
	Anexo I – <i>Cookies</i> utilizados pelo continente.pt	93

Índice de Figuras

Figura 1- Setores de atividade com aplicação dos Big Data e respetivos benefícios (Almeida, 2017).....	7
Figura 2- Os 6V's dos Big Data (Adaptado de: Demchenko, De Laat and Membrey, 2014)	8
Figura 3- Ferramenta de pesquisa fornecida pela Cookiepedia.....	36
Figura 4- Exemplo de <i>site</i> em que a Cookiepedia não consegue obter dados.....	36
Figura 5- Exemplo de pesquisa na Cookiepedia de um <i>cookie</i> específico.....	39
Figura 6- Sexo dos inquiridos	41
Figura 7- Tipo de Dispositivo que os inquiridos mais utilizam para aceder à Internet.	43
Figura 8- Dados que os inquiridos consideram como pessoais.....	44
Figura 9- Percentagem de inquiridos que costumam fornecer os seus dados pessoais em contexto de comércio eletrónico.....	44
Figura 10- Principais razões para fornecer dados pessoais	45
Figura 11- Nível de controlo sobre a informação disponibilizada online	45
Figura 12- Considerar aceitável opções que os <i>sites</i> oferecem para a sua utilização....	46
Figura 13- Conhecimento do que é o Regulamento Geral de Proteção de Dados	47
Figura 14- Opção correta para um conjunto de questões sobre o RGPD.....	47
Figura 15- Nível de concordância dos inquiridos sobre as condições que os <i>sites</i> oferecem para a sua utilização.....	48

Índice de Tabelas

Tabela 1- Resumo das principais tecnologias dos Big Data.....	11
Tabela 2- Benefícios do uso dos Big Data.....	13
Tabela 3- Desafios do uso dos Big Data.....	14
Tabela 4- Oportunidades e riscos da aplicação dos Big Data.....	15
Tabela 5- Informação pessoal recolhida por <i>sites</i> de comércio eletrónico	18
Tabela 6- Mecanismos de recolha de informações dos consumidores <i>online</i>	19
Tabela 7- Objetivos dos <i>Cookies</i>	25
Tabela 8- Correspondência: Objetivos e Questões do Questionário	34
Tabela 9- <i>Sites</i> selecionados no estudo de <i>Cookies</i>	37
Tabela 10- Faixa etária dos inquiridos	41
Tabela 11- Grau de Escolaridade dos Inquiridos.....	42
Tabela 12- Opinião dos inquiridos acerca de 5 afirmações relacionadas com os seus sentimentos relativamente ao terem de fornecer informações pessoais online.	49
Tabela 13- Tipos de <i>Cookies</i> utilizados pelos <i>sites</i> de comércio eletrónico selecionados	65
Tabela 14- Objetivo dos <i>Cookies</i> utilizados pelos <i>sites</i> de comércio eletrónico selecionados.....	66
Tabela 15- Principais <i>cookies</i> utilizados pelo continente.pt.....	67
Tabela 16- Exemplos de <i>cookies</i> utilizados pela decathlon.com.....	68
Tabela 17- Exemplos de <i>cookies</i> utilizados pela fnac.pt	69
Tabela 18- Exemplos de <i>cookies</i> utilizados pela sportzone.pt	70
Tabela 19- <i>Cookies</i> utilizados pela wook.pt	71
Tabela 20- <i>Cookies</i> utilizados em Worten.pt.....	72
Tabela 21- <i>Cookies</i> utilizados em custojusto.pt	73

Lista de abreviaturas

GPS - Global Positioning System (Sistema de Posicionamento Global)

IMPIC - Instituto dos Mercados Públicos, do Imobiliário e da Construção

IPI - Informações Pessoais Identificáveis

OCDE - Organização para a Cooperação e Desenvolvimento Económico

RFID - Radio Frequency Identification (Identificação por Radiofrequência)

RGPD - Regulamento Geral de Proteção de Dados

UE – União Europeia

CAPÍTULO I - INTRODUÇÃO

A evolução tecnológica nas últimas décadas aumentou significativamente o volume de dados disponíveis para as organizações, tendo impacto nos seus processos decisórios e resultados. Sistemas relacionados com os *Big Data* têm adquirido cada vez mais importância nos negócios nos últimos anos, visto que as organizações necessitam, cada vez, mais de analisar dados abundantes e variados (Chen, Chiang & Storey, 2012).

A euforia em torno dos *Big Data* foi indiscutivelmente gerada acima de tudo pela *web* e pelas comunidades presentes no comércio eletrónico. As transformações mais significativas foram alcançadas pelos principais fornecedores de comércio eletrónico, como a Amazon e o Ebay, através das suas plataformas inovadoras e pelos sistemas de recomendação de produtos.

As principais organizações da *Internet*, como a Google, a Amazon e o Facebook, continuam a liderar o desenvolvimento das plataformas *web*, utilizando *cloud computing*, redes sociais e plataformas de comércio eletrónico. O surgimento dos conteúdos da *web* 2.0 gerados pelo consumidor em vários fóruns, grupos de notícias, plataformas de redes sociais e sistemas de *crowdsourcing* oferece outra oportunidade para investigadores e profissionais “ouvirem” a voz do mercado, mercados esses que incluem clientes, funcionários, investidores e media (Chen, Chiang & Storey, 2012).

A motivação para o estudo apresentado nesta dissertação, deve-se à falta de estudos relacionados com esta área. A realização da investigação justifica-se assim pela existência de uma lacuna no estudo dos *Big Data* aplicados ao comércio eletrónico. Este tipo de estudo é bastante importante para uma organização, visto que os *Big Data* representam para essa organização um grande potencial de valor.

Este estudo contribui para os utilizadores *online* terem a perceção real de qual a utilização dada aos seus dados pessoais recolhidos em ambiente de comércio eletrónico, e para que as organizações adquiriram conhecimento relativamente às preocupações dos utilizadores *online* quanto aos seus dados pessoais.

O objetivo principal desta dissertação foi o de analisar em que medida os *Big Data* estão presentes e são utilizados no contexto de comércio eletrónico, quer do ponto de vista dos utilizadores, quer do ponto de vista das organizações. Para atingir o objetivo principal deste estudo, foram traçados os seguintes objetivos específicos:

- Perceber em que medida os Big Data estão presentes nas organizações de comércio eletrónico
- Identificar as características e atributos dos Big Data
- Enunciar as tecnologias mais presentes no âmbito do estudo dos Big Data
- Definir os benefícios e desafios, oportunidades e riscos dos Big Data
- Definir as questões legais e éticas da utilização dos dados do utilizador *online* em ambiente de comércio eletrónico
- Analisar qual a informação recolhida e os métodos de recolha pelos *sites* em contexto de comércio eletrónico
- Avaliar o nível de consciencialização do utilizador em relação à partilha e recolha dos seus dados
- Definir quais os dados recolhidos pelos *sites* de comércio eletrónico a partir de uma análise aos *cookies* existentes num conjunto selecionado de *sites* de comércio eletrónico.

Este estudo classifica-se, de acordo com os seus procedimentos, como um *Estudo de Caso*. A abordagem é *quantitativo-qualitativa* do tipo *descritiva*. A recolha de dados foi por meio de um questionário *online* cujo a amostra é *não-probabilística*, do tipo *bola de neve*, e uma análise a *cookies* de um conjunto de *sites*, com uma amostra por *conveniência*.

Foram elaborados dois instrumentos de recolha de dados: i) um questionário *online* para recolha de dados primários, de modo a perceber o ponto de vista do utilizador; e ii) a ferramenta de pesquisa disponibilizada em www.cookiepedia.com, para uma análise aos *cookies* de um determinado conjunto de *sites*, de modo a perceber o ponto de vista das organizações.

A partir das 160 respostas obtidas no questionário, concluiu-se que os inquiridos não têm noção de quais são de facto os dados pessoais que partilham e que apesar de os inquiridos sentirem pouco ou nenhum controlo sobre a informação que disponibilizam, têm o hábito de aceitar os *cookies* quando acedem a um *site*, que não utilizam software que evite a monitorização da sua atividade nem que mudam as definições de privacidade do navegador que utilizam. No entanto, grande parte dos inquiridos evita aceder a alguns *sites* de comércio eletrónico com receio que a sua atividade seja monitorizada. Foi também possível relacionar as Políticas de Privacidade de todos os *sites* analisados e os

resultados dos *cookies* utilizados e perceber que a publicidade *online* é a utilização mais comum dos chamados *Third Party Cookies*. A grande maioria dos *sites* analisados partilha os dados recolhidos com empresas terceiras. São poucos os *sites* que utilizam *cookies* como análise de dados ou funcionalidade do *site* e que os *cookies* estritamente necessários têm pouca presença nos *cookies* utilizados pelos *sites* de comércio eletrónico.

O restante deste documento é composto por seis capítulos, encontrando-se estruturado da seguinte forma:

No segundo capítulo, *Referencial Teórico*, é apresentada um breve suporte teórico que faculta conhecimentos relativos à área dos Big Data, incluindo conceitos recentes de Big Data e os seus seis atributos mais significativos. São apresentados breves exemplos da aplicação dos Big Data e são apresentadas as três tecnologias mais relevantes: MapReduce, Hadoop e NoSQL. É também feito um breve resumo dos benefícios, desafios, oportunidades e riscos dos Big Data. Expõem-se algumas questões éticas e legais no contexto dos Big Data, onde se refere a importância da proteção de dados nos Big Data e no comércio eletrónico e se menciona o Regulamento Geral de Proteção de Dados Pessoais na Europa. Explica-se também o conceito de *online tracking* onde se menciona a utilidade dos *cookies* e os seus objetivos.

No capítulo seguinte, *Metodologia*, apresenta-se a abordagem metodológica e os métodos adotados para a condução da pesquisa. Apresenta-se a abordagem que foi feita à investigação e os instrumentos utilizados para recolha de dados.

Em seguida, no quarto capítulo, *Resultados*, apresentamos os dados recolhidos e trabalhados do questionário *online* e do estudo relativo à análise às Políticas de Privacidade anunciadas e aos *cookies* utilizados por *sites* de comércio eletrónico.

No quinto capítulo, *Discussão*, estabelecem-se as relações entre os diversos resultados e suas implicações gerais, através de uma triangulação dos dados, a partir das quais se irão retirar conclusões.

Por fim, no sexto e último capítulo, é apresentada a *Conclusão*, a partir dos objetivos propostos inicialmente. Para além disso, também são apresentadas algumas limitações encontradas na investigação realizada e são apontadas sugestões para investigação futura.

CAPÍTULO II – REFERENCIAL TEÓRICO

Neste capítulo, é apresentado o referencial teórico, com base em estudos de diversos autores acerca de temas como o conceito de Big Data, os seus V's, as principais tecnologias de aplicação, abordando por fim os benefícios e as limitações dos Big Data. São também abordadas algumas questões éticas e legais.

2.1 Conceito de Big Data

Raheem (2019, p.3) refere-se ao termo “Big Data” como um grande volume de dados – sejam estruturados (por exemplo, uma base de dados relacional), sejam semiestruturados (arquivo em XML), ou sejam não estruturados (emails, ficheiros de texto, vídeos...) que abundam nas organizações no dia-a-dia. Sharda, Delen & Turban (2017, p.401) acrescentam ainda que o termo “Big Data” tem sido usado para descrever volumes massivos de dados analisados por organizações em grande escala, tais como a Google, e projetos de pesquisa científica da NASA. Para a maioria das organizações, é um termo relativo: “Big” depende do tamanho de uma organização. O foco é mais sobre encontrar novo valor dentro e fora das fontes de dados convencionais. Descrevem ainda que os Big Data excedem o alcance de ambientes de *hardware* comumente usados ou a capacidade das ferramentas de *software* para capturar, gerir e processar todos os dados dentro de um intervalo de tempo tolerável para os seus utilizadores. No entanto, não é importante apenas a quantidade de dados, mas sim o que é que as organizações fazem com esses dados. Os Big Data podem ser analisados na procura de *insights* que levem a melhores decisões e movimentos estratégicos dos negócios. São essas informações valiosas que impulsionam a inteligência de negócios, que é tão importante para o seu crescimento.

É necessário questionar-se de onde é que os Big Data vêm. A resposta é: de todo o lado (Sharda, Delen & Turban, 2017, p.401). Entre as suas fontes crescentes estão o *streaming*, as sequências de *clicks* da *Web*, os conteúdos das redes sociais (*tweets*, *blogs*, *posts* no Facebook), os vídeos, entre outras formas de entretenimento.

Estas fontes de dados, que anteriormente foram ignoradas por causa das limitações técnicas, são agora tratadas como minas de ouro. Big Data abrange desde os dados de voz de um *call center* até aos dados genómicos e proteómicos da pesquisa biológica e da medicina (Davenport, Barth & Bean, 2012). Pode surgir de sensores RFID¹, sistemas de

¹ Em português, Identificação por Radiofrequência, é um método de identificação automática através de sinais de rádio, recuperando e armazenando dados remotamente.

GPS², sensores em tempo real, redes sociais, índices de pesquisa na *Internet*, registros de chamadas, astronomia, vigilância militar, pesquisa científica, arquivos fotográficos e da prática de comércio eletrônico em grande escala.

Almeida (2017) apresenta cinco exemplos de setores de atividade nos quais os Big Data são utilizados com benefícios, nomeadamente o da Tecnologia de Informação, o do Atendimento ao cliente, o do Comércio eletrônico, o dos Bancos e Mercados Financeiros. Estes são apresentados na Figura 1.

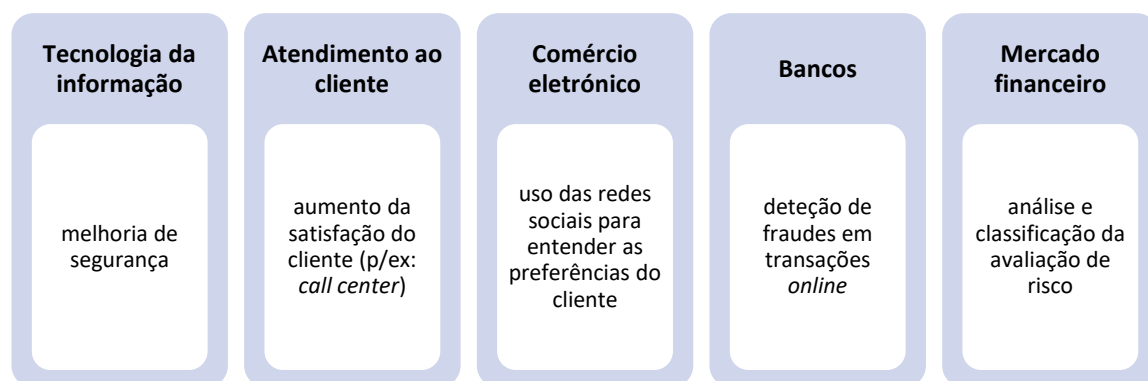


Figura 1- Setores de atividade com aplicação dos Big Data e respectivos benefícios (Almeida, 2017)

A adoção dos Big Data nas organizações, independentemente do seu tamanho, tem sido atualmente um dos maiores desafios para a sua implementação. Ainda há muitas organizações que lutam para entrar nesse novo mundo da informação e outras que já usam esta tecnologia, mas ainda de forma limitada e restrita. A integração, manipulação e qualidade dos Big Data surgem como pontos-chave que devem ser considerados na construção de uma solução de gestão dos Big Data (Kaur; Monga, 2016).

2.2 Os 6 V's dos Big Data

A origem do conceito de Big Data remonta à década de 1990. O termo foi usado pela primeira vez por John Mashey, ex-cientista chefe aposentado da Silicon Graphics, para se referir ao manuseio e análise de conjuntos massivos de dados (Diebold, 2012). Doug Laney (2001) detalhou que os Big Data eram caracterizados por três atributos: **Volume**, **Velocidade** e **Variedade**. Estes ficaram conhecidos como os 3 V's do Big Data.

² Em português, Sistema de Posicionamento Global, um sistema de navegação por satélite que fornece a um aparelho receptor móvel a sua posição a qualquer momento e em qualquer lugar na Terra; desde que o receptor se encontre no campo de visão de três satélites GPS.

Taurion (2013) acrescenta mais dois atributos, sendo eles **Veracidade** e **proposta de Valor**, sendo este último atributo bastante importante na análise da presença dos Big Data nas redes sociais. Mais tarde, Sharda, Delen & Turban (2017) referem também a existência de outro atributo, a **Variabilidade**. Na Figura 2 apresenta-se um resumo dos V's que serão abordados ao longo desta secção. Apesar de neste estudo apenas serem abordados 6 V's dos Big Data, é de salientar a existência de bastantes mais atributos, como a **Validade**, **Volatibilidade**, **Vulnerabilidade**, entre outros (Kitchin; McArdle, 2016).

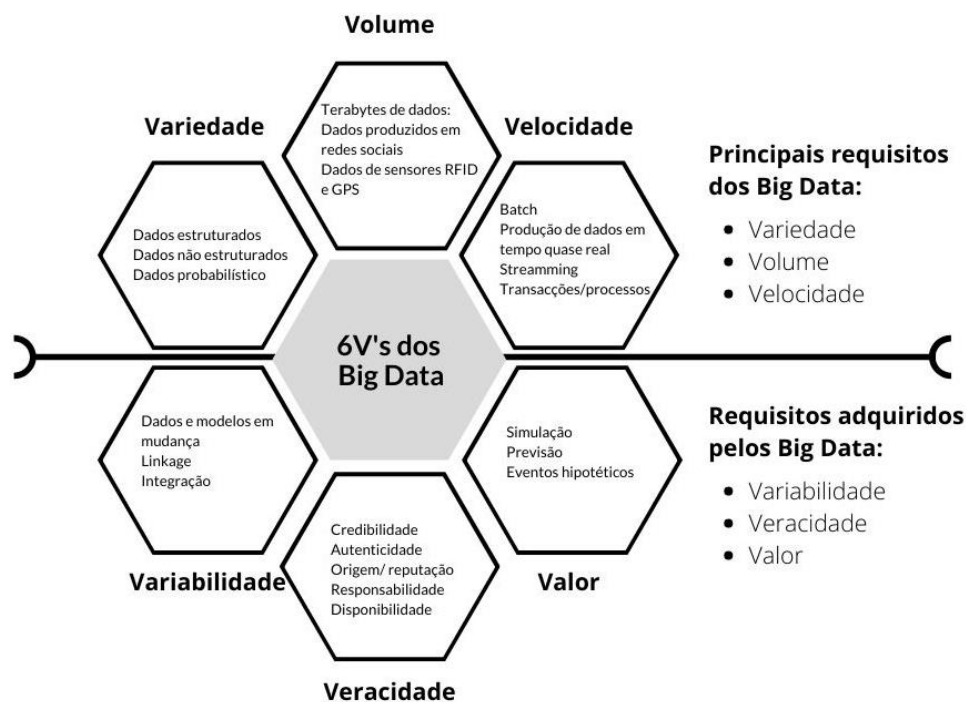


Figura 2- Os 6V's dos Big Data (Adaptado de: Demchenko, De Laat and Membrey, 2014)

“O **volume** está relacionado com a quantidade crescente de dados, que impactam diretamente processos organizacionais e influenciam métodos preditivos e estatísticos.” (Félix, Tavares & Cavalcante, 2018, p.113). Muitos fatores contribuíram para o aumento exponencial do volume de dados, tais como dados de sistemas de processamento de transações gerados ao longo dos anos, dados de texto continuamente transmitidos nas redes sociais, quantidades crescentes de dados de sensores RFID e GPS, os quais são constantemente gerados de forma automática (Sharda, Delen & Turban, 2017).

A quantidade de informação originada das redes sociais, por exemplo, é tão antiga quanto as próprias redes. O Facebook, lançado em 2004 e o Twitter, em 2006, e a recente proliferação de *smartphones*, *tablets*, redes *wireless* e redes de dados 3G, 4G e 5G,

permitiu aos utilizadores estarem conectados praticamente em qualquer momento, gerando conteúdos, ideias e opiniões. Como esses dispositivos são onnipresentes, é fácil esquecer que o iPhone foi revelado apenas em 2008 e o iPad apenas em 2010. Assim, as bases de dados estruturados que armazenaram a maioria da informação das organizações, até recentemente, são inadequadas para armazenar e processar grandes volumes de dados, normalmente não estruturados ou semiestruturados. Ao mesmo tempo, o constante declínio dos custos de todos os elementos da computação - armazenamento, memória, processamento, largura de banda e assim por diante - significa que as abordagens intensivas de dados, que antes representavam custos demasiado elevados, estão a tornar-se rapidamente mais económicas (McAfee & Brynjolfsson, 2012).

A **Variedade** define os vários tipos de formatos que são gerados, sejam eles dados estruturados, tais como dados numéricos em bases de dados tradicionais, documentos de texto não estruturados, e-mail, vídeos, áudio, *likes e dislikes* nas redes sociais, coordenadas de GPS, dados de sensores RFID, *upload* de fotos, filmagens transmitidas em tempo real ou até dados de cotações da bolsa e transações financeiras (Machado, 2018). Segundo algumas estimativas, 80 a 85% de todos os dados das organizações estão sem algum tipo de estrutura ou em formato semiestruturado. Mas não há como negar o seu valor e, portanto, devem ser incluídos nas análises para apoio à tomada de decisão (Sharda, Delen & Turban, 2017). O grande desafio dos Big Data está precisamente em lidar com diferentes formatos de dados e informação, todos os dias, simultaneamente.

A **Velocidade** refere-se à velocidade crescente a que estes dados são criados, ou seja, a velocidade crescente a que os dados podem ser processados, armazenados e analisados por bases de dados relacionais (Ishwarappa, Anuradha, 2015). Os dados produzidos por sensores RFID, dispositivos de GPS e medidores inteligentes estão a impulsionar a necessidade crescente de lidar com quantidades massivas de dados quase em tempo real. Reagir com rapidez suficiente para lidar com a velocidade é o grande desafio para as organizações (Sharda, Delen & Turban, 2017). McAfee & Brynjolfsson (2012) afirmam que há mais dados que se cruzam na *Internet* a cada segundo do que aqueles que foram armazenados em toda a *Internet* em 20 anos. Isso dá às organizações uma oportunidade de aceder a muitos *petabytes*³ de dados num único conjunto de dados - e não apenas na *Internet*. Por exemplo, estima-se que o Walmart recolhe e armazena mais de 2,5 *petabytes*

³ Um petabyte é um quatrilhão de bytes, ou o equivalente a cerca de 20 milhões de arquivos de armários. Um exabyte é 1.000 vezes esse valor, ou um bilhão de gigabytes.

de dados a cada hora nas suas transações com os clientes, que a cada minuto de cada dia são carregados centenas de horas de vídeo no Youtube e que são enviados mais de 200 milhões de *emails* através do Gmail (Ishwarappa, Anuradha, 2015).

A **Veracidade**, de acordo com Sharda, Delen & Turban (2017), é um termo atribuído pela IBM para definir o quarto “V” para descrever Big Data. Este atributo refere-se à conformidade dos factos: precisão, qualidade ou confiança nos dados (Sharda, Delen & Turban, 2017). Quando lidamos com alto volume, velocidade e variedade de dados, é quase impossível que todos os dados estejam 100% corretos, o que faz com que a qualidade dos dados recolhidos possa variar muito. A precisão da análise dos dados depende da veracidade da fonte dos dados (Ishwarappa, Anuradha, 2015). Existem ferramentas e técnicas que são usadas para vincular a veracidade dos Big Data, transformando os dados em *insights* de qualidade e confiáveis (Sharda, Delen & Turban, 2017).

A **Variabilidade** considera as instabilidades do fluxo de dados, devido às velocidades crescentes e variedades de dados (Sharda, Delen & Turban, 2017). Por exemplo, o repositório do Twitter armazena e gere diferentes tipos de dados. Estes dados podem ter vários sentidos e significados, uma vez que estamos a lidar com uma rápida mudança de fluxo de dados, visto que, em quase todos os *tweets*, a mesma palavra pode ter um significado totalmente diferente (Sivarajah et al, 2016). Por exemplo, para realizar uma análise de sentimentos adequada, Zhang, Hu et al. (2015) afirmam que os algoritmos precisam ser capazes de entender o contexto e ser capazes de decifrar o significado exato de uma palavra nesse contexto.

A **proposta de Valor** é o aspeto mais importante nos Big Data. É bom ter acesso a Big Data, mas a menos que possamos transformá-los em valor, eles tornam-se inúteis. (Ishwarappa, Anuradha, 2015). Ao analisar Big Data, as organizações podem obter mais valor de negócio que de outra forma não conseguiriam. Enquanto que os utilizadores podem detetar padrões em pequenos conjuntos de dados usando métodos simples de estatística e *machine learning* ou ferramentas simples de consulta e relatórios, os Big Data permitem a análise de dados em grande escala, o que significa maior *insight* e melhores decisões (Sharda, Delen & Turban, 2017).

2.3 Tecnologias no âmbito dos Big Data

Big Data é um termo bastante amplo e a diversidade de tecnologias e produtos existentes é igualmente ampla, surgindo novidades regularmente. Há várias tecnologias para processar e analisar Big Data. No entanto, iremos apenas abordar as três tecnologias de Big Data que mais se destacam, o **Apache Hadoop**, **MapReduce** e **NoSQL** (Sharda, Delen & Turban, 2017). Apresenta-se na Tabela 1 um resumo destas tecnologias com mais valor na área dos Big Data.

Tabela 1- Resumo das principais tecnologias dos Big Data

Capacidade dos Big Data	Tecnologia Primária	Características
Base de dados	NoSQL	Desenho dinâmico e flexível de esquemas, altamente escalável, centro de dados múltiplos, tolerante a falhas, bases de dados de alto desempenho de pares de valores-chave
Processamento	MapReduce	Distribuição da carga de trabalho de dados por milhares de nós. Quebra o problema em sub-problemas menores
	Apache Hadoop	Infraestrutura altamente personalizável, Processamento em lote paralelo altamente escalável, tolerante a falhas

Adaptado de Abhay & Dhanya (2017)

2.3.1 Apache Hadoop

O Apache Hadoop é uma infraestrutura que permite o processamento distribuído de grandes conjuntos de dados através de *clusters* de computadores usando modelos de programação simples. É projetado para escalar de servidores individuais para milhares de servidores, cada um oferecendo computação e armazenamento local. Em vez de confiar no *hardware* para fornecer alta disponibilidade, a própria biblioteca é projetada para detetar e lidar com falhas na camada de aplicação, fornecendo assim um serviço altamente disponível em cima de um *cluster* de computadores, cada um dos quais pode ser propenso a falhas (Apache Hadoop, 2019).

2.3.2 MapReduce

É uma técnica popularizada pela Google que distribui o processamento de ficheiros multi-estruturados de grande dimensão por um grande aglomerado de máquinas. O seu maior desempenho é obtido dividindo o processamento em pequenas unidades de trabalho que podem ser executadas em paralelo em centenas, ou até milhares de nós no aglomerado de máquinas (Sharda, Delen & Turban, 2017).

2.3.3 NoSQL

O **Not Only SQL** (NoSQL) refere-se a um grupo eclético e cada vez mais familiar de sistemas de gestão de bases de dados não-relacionais, as quais não são construídas baseando-se em estruturas que são tabelas, e geralmente não utilizam a linguagem SQL para manipulação de dados. Estes sistemas são úteis quando se trabalha com uma grande quantidade de dados cuja natureza não exige um modelo relacional (Moniruzzaman; Hossain, 2013). Os sistemas de base de dados NoSQL surgiram em conjunto com grandes empresas de *Internet*, tais como Google, Amazon e Facebook, que tinham desafios para lidar com enormes quantidades de dados (Moniruzzaman; Hossain, 2013).

2.4 Benefícios e desafios, oportunidades e riscos dos Big Data

Devido ao seu grande valor, os Big Data têm essencialmente mudado e transformado a nossa forma de viver, trabalhar e pensar. Estamos, atualmente, na era da informação e o uso extensivo da *Internet*, *Internet das Coisas*, do *Cloud Computing*, e de outras tecnologias de informação emergentes, têm feito várias fontes de dados aumentar a um ritmo sem precedentes, tornando as estruturas e tipos de dados cada vez mais complexos (Bhadani; Jothimani, 2016). Na Tabela 2 apresentamos um resumo dos benefícios do uso dos Big Data segundo alguns autores.

Tabela 2- Benefícios do uso dos Big Data

Benefícios do uso dos Big Data	Autores
<ul style="list-style-type: none">• “Captura e análise de grandes volumes de dados estruturados e não estruturados• Melhoria das capacidades analíticas gerais• Análise de informação em tempo real• Redução de custos da análise e processamento de dados• Maior transparência das informações utilizadas para a tomada de decisão• Personalização mais ágil de produtos e serviços• Criação de novos modelos de negócios”	(Félix, Tavares & Cavalcante (2018, p. 113).
<ul style="list-style-type: none">• Redução de custos (as ferramentas dos Big Data permitem que as organizações armazenem grandes volumes de dados a um preço muito mais baixo do que uma base de dados tradicional)• Vantagem competitiva para as organizações, oferecendo a possibilidade de explorar novas oportunidades de negócios	(Almeida, 2017).
<ul style="list-style-type: none">• Aumentar a percepção do comportamento do consumidor;• Aumento de vendas;• Aumento de inscrições e registros;• Aumento do retorno do investimento (ROI);• Aumentar a satisfação do cliente;• Aumentar as oportunidades de vendas.	Porres (2013)
<ul style="list-style-type: none">• “Redução dos custos e aumento das receitas;• Aumento da eficiência operacional;• Melhoria na tomada de decisão;• Melhoria de produtos e serviços;• Melhoria nos processos de inovação e de desenvolvimento de novos produtos e mercados.”	Félix, Tavares & Cavalcante (2018, p. 114)

No entanto, existem atualmente muitos desafios relacionados com o aproveitamento do potencial dos Big Data, sendo alguns causados pelas suas próprias características, outros pelos modelos e métodos de análise atuais e outros pelas limitações dos sistemas de processamento de dados atuais (Jin *et al.*, 2015). A Tabela 3 apresenta alguns desafios na utilização dos Big Data, segundo dois autores.

Tabela 3- Desafios do uso dos Big Data

Desafios do uso dos Big Data	Autores
<ul style="list-style-type: none">• Armazenamento e análise de dados;• Descoberta de conhecimento e complexidades computacionais;• Escalabilidade e visualização de dados;• Segurança da informação.• Os dispositivos de recolha de dados são cada vez mais variados, desde dispositivos móveis, sensores e redes sociais que produzem enormes quantidades de dados em mudança. Para obter o maior valor possível a partir desses dados, as organizações precisam de processar dados e tomar decisões muito mais rapidamente	Almeida (2017)
<ul style="list-style-type: none">• O armazenamento disponível não é suficiente para armazenar a grande quantidade de dados que está a ser produzida (os <i>sites</i> de redes sociais são, eles próprios, um grande contribuinte juntamente com os dispositivos sensores).	Akter e Wamba (2016)

Por causa das rigorosas exigências dos Big Data, o armazenamento dos dados em servidores terceiros e na nuvem pode parecer uma opção. No entanto, carregar esta grande quantidade de dados na nuvem não resolve o problema. É necessária a obtenção de todos os dados recolhidos e, em seguida, ligá-los numa forma que permita extrair as informações importantes. Ou seja, *Terabytes* de dados iriam demorar muito tempo a serem carregados na nuvem e, além disso, esses dados mudam tão rapidamente que isso fará com que os dados sejam difíceis de serem carregados em tempo real (Akter; Wamba, 2016). Lenz (2019), nomeia algumas oportunidades e riscos do uso dos Big Data, como podemos observar na Tabela 4.

Tabela 4- Oportunidades e riscos da aplicação dos Big Data

Oportunidades dos Big Data	Riscos dos Big Data
Aumenta a transparência e a informação	Invasão de privacidade humana
Análise preditiva e <i>machine learning</i>	Rastreamento e perfilagem de indivíduos
Automação da decisão de processos	Ataques informáticos, fraude e roubo de identidade
Aumenta a eficácia no uso dos recursos	Informações falsas, notícias, vídeos
Conhecimento das preferências dos clientes	Expansão das assimetrias de poder e da desigualdade social
Estratégias de <i>marketing</i> personalizadas	Discriminação de serviços e preços, risco de exclusão

Adaptado de Lenz (2019)

A comparação dos riscos e oportunidades acima mostra claramente a necessidade de regulamentação adequada dos Big Data para proteger os direitos individuais à privacidade e à não-discriminação. Isto aplica-se especialmente a grupos economicamente mais fracos na sociedade (Lenz, 2019). Esta questão da privacidade e partilha de dados será desenvolvida ao longo da secção 2.5, onde serão abordadas as questões éticas e legais relacionadas com esta temática.

2.5 Questões éticas e legais no contexto dos Big Data

A *Internet* e a sua utilização têm suscitado uma crescente preocupação ética, social e política numa escala sem precedentes. A *Internet* é a raiz de diversas controvérsias contemporâneas e parte da resposta a estas controvérsias reside nas características subjacentes à própria tecnologia da *internet* e nas formas como ela tem sido explorada pelas empresas.

Muitas empresas e indivíduos beneficiam do desenvolvimento comercial da *Internet*, mas este desenvolvimento também cobra um preço aos indivíduos, organizações e sociedade. Estes custos e benefícios devem ser cuidadosamente considerados por aqueles que procuram tomar decisões ética e socialmente responsáveis neste novo ambiente (Laudon & Traver, 2019). A *internet*, em geral, e o comércio eletrónico em particular, levantam tantas questões éticas, sociais e políticas que é difícil classificá-las todas, sendo assim

complicado ver o seu relacionamento umas com as outras. Claramente que todas estas questões estão interligadas.

Katal, Wazid & Goudar (2013) consideram que os grandes desafios dos Big Data são a privacidade, segurança, o acesso aos dados e a partilha e recolha de informação. Referem ainda que a questão da privacidade e segurança é bastante sensível e inclui um significado conceptual, técnico e jurídico, visto que:

- As informações pessoais de uma pessoa quando combinadas com grandes conjuntos de dados externos leva à inferência de novas informações sobre essa pessoa e é possível que essas informações sejam sigilosas e a pessoa pode não querer que o proprietário dos dados ou que qualquer pessoa saiba sobre eles.
- A informação relativa aos utilizadores (pessoas) é recolhida e utilizada de forma a acrescentar valor ao negócio da organização. Isto é feito através da criação de perspetivas sobre as suas vidas das quais eles desconhecem.
- Os Big Data utilizados pelas autoridades policiais aumentarão as chances de certas pessoas marcadas sofrerem consequências adversas sem a capacidade de reagir ou mesmo ter conhecimento de que estão a ser discriminadas.

2.5.1 Proteção de dados nos Big Data e no comércio eletrónico

Quando os primeiros sistemas informatizados de grande escala e de âmbito nacional começaram a aparecer nos Estados Unidos, na década de 1960, surgiram questões relacionadas com a privacidade. Por exemplo, os sistemas de cartão de crédito, pela primeira vez, deram aos comerciantes a capacidade de recolher informação em formato digital sobre o comportamento do consumidor. Pela primeira vez, havia bases de dados nacionais privadas que continham informações sobre o que as pessoas compraram, a quem e onde (Laudon & Traver, 2019). Outras instituições dos setores de educação, saúde e serviços financeiros também começaram a criar bases de dados de grande escala que envolviam milhões de cidadãos. Estes desenvolvimentos levaram aos primeiros esforços para reivindicar o direito à privacidade do consumidor.

No entanto, ninguém antecipou a ascensão da Google, Facebook, Amazon, Netflix e outras empresas de comércio eletrónico que recolhem informações pessoais sobre quase toda a população dos Estados Unidos e do mundo (Laudon & Traver, 2019). Ninguém previu que o comércio eletrónico iria envolver mais de 200 milhões de pessoas só nos

Estados Unidos, que uma única empresa dominaria a venda de livros (Amazon) ou o aluguer de filmes (Netflix), que uma única empresa iria dominar o mercado da pesquisa *online* e recolher dados muito pormenorizados sobre as intenções dos consumidores e interesses (Google) ou que uma única empresa se tornaria um repositório da vida social de mil milhões de pessoas (Facebook) (Laudon & Traver, 2019). Isto é, ninguém previu o papel relevante dos Big Data no contexto do comércio eletrónico.

Os Big Data também aumentam a importância da harmonização, ou mesmo da normalização, nas normas de proteção de dados. Dado que os dados pessoais são universalmente recolhidos e partilhados através de fronteiras sectoriais e nacionais, a incoerência das leis de proteção de dados constitui uma ameaça crescente para os indivíduos, as instituições e a sociedade (Kuner *et al.*, 2012). Talvez o maior impacto dos Big Data seja a pressão que estes trazem para um novo debate ponderado, informado e multinacional sobre os princípios fundamentais que devem estar na base da proteção dos dados. A maior parte das leis de proteção de dados continua a basear-se nas Diretrizes da OCDE de 1980. Estas Diretrizes resistiram bem às três décadas que se seguiram, mas é importante recordar que foram elaboradas não apenas antes dos Big Data, mas também antes de a World Wide Web, computadores portáteis, GPS, *smartphones*, dispositivos tablet, ou a miríade de outras inovações que tornam possíveis os Big Data (Kuner *et al.*, 2012).

2.5.1.1 Informação recolhida pelos sites

Os *sites* de comércio eletrónico recolhem diariamente uma variedade de informação de ou sobre consumidores que visitam o seu *site* e/ou fazem compras. Alguns desses dados constituem Informações Pessoais Identificáveis (**PII**), que são definidas como quaisquer dados que possam ser utilizados para identificar, localizar ou contactar um indivíduo (Laudon & Traver, 2019). Outros dados são informações anónimas, onde a identidade da pessoa não é um nome, mas um código atribuído. Essas informações anónimas incluem informações demográficas e comportamentais, como idade, ocupação, rendimento, código postal, etnia e comportamento de navegação na *web*, sem identificar quem nós somos (Laudon & Traver, 2019).

A Tabela 5 lista algumas das **PII** recolhidas diariamente por *sites* de comércio eletrónico, incluindo *sites* e aplicações móveis. Embora anónima, esta informação detalhada ainda é pessoal e um identificador de nome pode ser facilmente anexado às informações. No

entanto, os anunciantes e comerciantes *online* estão menos interessados no nosso nome do que no nosso comportamento de compra.

Tabela 5- Informação pessoal recolhida por *sites* de comércio eletrónico

Informação pessoal recolhida por <i>sites</i> de comércio eletrónico		
Nome	Género	Educação
Morada	Idade	Preferências
Número de telemóvel	Ocupação	Transações
E-mail	Localização	Dados de <i>Clickstream</i>
Dados de contas bancárias	Gostos (<i>Likes</i>)	Tipo de navegador

Adaptado de Laudon & Traver (2019)

As redes de publicidade e mecanismos de procura também rastreiam o comportamento dos consumidores em milhares de *sites* populares e não apenas num único *site*, via *cookies*, *web beacons*, software de rastreamento, *spyware* e outras técnicas. Por exemplo, basta clicar num botão “*Like*” num *site*, para que o Facebook possa seguir os nossos movimentos na *Web* (Laudon & Traver, 2019).

2.5.1.2 Principais formas pelas quais as empresas *online* recolhem informação sobre os utilizadores

Existem vários mecanismos para a recolha de informação sobre os consumidores *online*. Na Tabela 6 são apresentados os principais mecanismos de recolha de informação e qual o seu impacto na privacidade dos utilizadores *online*. Sendo um dos pontos fulcrais deste estudo, apenas serão desenvolvidos os mecanismos de *Cookies* e *Third Party Cookies*, na secção 2.6.1.1.

Tabela 6- Mecanismos de recolha de informações dos consumidores *online*.

Mecanismos	Impacto Na Privacidade
<i>Smartphones e apps</i>	Usado para rastrear a localização e compartilhar fotos, endereços, números de telefone, e outros comportamentos.
Redes de Publicidade	Usado para rastrear indivíduos que se movimentam entre milhares de <i>sites</i> .
Redes Sociais	Usado para recolher informações sobre conteúdo fornecido pelo utilizador, como livros, música, amigos e outros interesses, preferências e estilos de vida.
<i>Cookies e Supercookies</i>	Usado para rastrear indivíduos num único local. <i>Supercookies</i> são quase impossíveis de identificar ou remover.
Spyware	Pode ser usado para registar toda a atividade de teclado de um utilizador, incluindo <i>sites</i> visitados e códigos de segurança usados; também usado para exibir anúncios a utilizadores com base nas suas pesquisas ou outro tipo de comportamento.
Search engine behavioural targeting⁴	Utiliza o histórico de pesquisas anteriores, dados demográficos, interesses expressos, geográficos ou outros dados inseridos pelo utilizador para segmentar a publicidade.
Carrinho de Compras	Pode ser usado para recolher informações detalhadas de pagamento e compra.
Formulários	Formulários <i>online</i> que os utilizadores preenchem voluntariamente em troca de um benefício ou recompensa prometidos que são vinculados ao <i>clickstream</i> ou a outros dados comportamentais para criar um perfil pessoal.
<i>Logs de transações do site</i>	Pode ser usado para recolher e analisar informações detalhadas sobre o conteúdo da página visualizada pelos utilizadores.
Motores de busca	Pode ser usado para rastrear declarações e opiniões que os utilizadores manifestam na <i>web</i> .

Adaptado de Laudon & Traver (2019)

Ao abordar as informações recolhidas pelos *sites* e as principais formas pelas quais as empresas *online* recolhem informação sobre os utilizadores, é fundamental abordar a

⁴ Usados pela Google e por outros motores de busca.

necessidade de regulação adequada, que garanta a privacidade dos utilizadores e defina o uso dos seus dados. Para isso, ao longo deste estudo, são abordadas algumas leis que obrigam o tratamento de dados de forma uniforme, nomeadamente o Regulamento Geral de Proteção de Dados (secção 2.5.2.) e a Lei n.º 46/2012, conhecida por “Lei dos *Cookies*” (secção 2.6.)

2.5.2 Regulamento Geral de Proteção de Dados Pessoais na Europa

De forma a responder aos novos desenvolvimentos tecnológicos e assegurar a livre circulação de dados pessoais no mercado, a União Europeia (UE) adotou em 2015 um regulamento com o objetivo de criar coerência entre as normas de proteção de dados em toda a UE. O Regulamento Geral de Proteção de Dados (RGPD) entrou em vigor a 25 de maio de 2018, trazendo algumas alterações significativas ao sistema de proteção de dados. Assim, o RGPD vem criar mais obrigações para as empresas, visto que estas devem aplicar as medidas técnicas adequadas para garantir o anonimato dos utilizadores *online*; manter os registos que provem a conformidade de todas as obrigações decorrentes do regulamento e, em alguns casos, preparar avaliações de impacto na privacidade e nomear responsáveis pela proteção de dados, sendo que estes devem fornecer informações sobre os dados pessoais que estão na posse das pessoas singulares. Qualquer violação de dados que possa causar danos aos direitos dos indivíduos deve ser notificado no prazo de 72 horas após a empresa tomar conhecimento do facto (Artigo 33º do RGPD da UE). O RGPD aplica-se a empresas ou entidades cujo tratamento de dados pessoais seja efetuado numa sucursal estabelecida na UE ou a empresas fora da UE que ofereçam serviços e bens ou controla os dados pessoais dos utilizadores na UE.

Os cidadãos veem consagrados novos direitos no RGPD da UE para proteger os seus dados pessoais. Dois dos mais importantes são o Direito à Portabilidade de Dados e o Direito a Ser Esquecido. Uma nota introdutória que se aplica a estes dois direitos é que as empresas, responsáveis pelo tratamento de dados, devem informar os titulares dos dados sobre a existência desses direitos e a possibilidade de os exercer, solicitando o apagamento dos dados ou a sua transferência para outro responsável pelo tratamento (Artigo 13º(2)(b) e Artigo 14º(2)(c) do RGPD da UE).

2.5.2.1 Direito à Portabilidade dos Dados

No que respeita ao direito à portabilidade dos dados, o artigo 20.º do RGPD da UE cria um direito totalmente novo que permite aos titulares "receber os dados pessoais que lhe

digam respeito e que tenha fornecido a um responsável pelo tratamento, num formato estruturado, de uso corrente e de leitura automática, e o direito de transmitir esses dados a outro responsável pelo tratamento a quem os dados pessoais foram fornecidos" (Artigo 20 (1) do RGPD da UE). O objetivo subjacente ao direito à portabilidade dos dados é capacitar as pessoas no que respeita à forma como podem controlar e gerir os seus próprios dados, criando meios para transmitir dados pessoais de um ambiente digital para outro e levando possivelmente a mais inovação e mais concorrência. (European Commission, 2017). Inclui dois elementos: o direito de receber dados pessoais de um responsável pelo tratamento de dados e de os conservar como entender; o direito de transmitir dados pessoais de um responsável pelo tratamento de dados para outro (European Commission, 2017).

Quanto ao tipo de dados que devem ser incluídos neste direito à portabilidade dos dados, o Artigo 20 (1) do RGPD da UE estabelece que devem ser dados que o sujeito deu consentimento ou que se baseiam no contrato por ele assinado. Devem ser dados relativos à pessoa em causa ou fornecidos por esta ao responsável pelo tratamento de dados.

O direito à portabilidade dos dados não pode afetar negativamente os direitos e liberdades de terceiros e tem de respeitar os direitos de propriedade intelectual e os segredos comerciais (European Commission, 2017). Aplica-se apenas aos dados “de forma automática”, o que significa que não inclui ficheiros em formato papel.

2.5.2.2 Direito a Ser Esquecido

O RGPD da UE inclui no artigo 17.º o Direito a Ser Esquecido, que refere que uma pessoa pode solicitar ao responsável pelo tratamento de dados que apague os seus dados pessoais. Este direito já tinha sido objeto de um acórdão do Tribunal de Justiça Europeu de 2014, após um cidadão espanhol ter apresentado uma queixa contra a Google por manter *online* páginas de jornais nos resultados de procura referentes a uma venda forçada do seu imóvel para saldar as suas dívidas, afetando assim a sua reputação (European Commission, 2016). O tribunal decidiu a seu favor e a Google foi forçada a cumprir com a decisão tendo aceiteado várias outras solicitações para remover URLs desde então.

Para exercer o direito de ser esquecido, devem estar presentes determinadas circunstâncias, nomeadamente que os dados pessoais já não sejam necessários para os fins originais de processamento, que o titular dos dados retire o seu consentimento para o

processamento de dados ou que o processamento de dados não seja lícito (Artigo 17 (1) do RGPD da UE).

Se os responsáveis pelo tratamento de dados tiverem tornado públicos os dados pessoais, podem ser obrigados a tomar as medidas necessárias para informar qualquer outra pessoa que esteja a tratar os dados que a pessoa em causa tenha solicitado para serem eliminados (Artigo 17 (2) do RGPD da UE).

O direito de apagar dados não deve ser visto como um direito absoluto. Ou seja, não se destina a ser usado para tornar as pessoas proeminentes menos proeminentes ou tornar os criminosos menos criminosos (European Commission, 2016). Deve ser equilibrado com o cumprimento das obrigações legais, outros interesses públicos e a proteção da liberdade de expressão e da liberdade dos meios de comunicação social.

2.6 Rastreamento Online

Desde o início da *web*, o número de terceiros que rastreiam e registam a atividade dos utilizadores explodiu. Alguns rastreadores têm uma visão substancial das atividades dos utilizadores em toda a *web*: a Google, por exemplo, tem uma presença de rastreio em cerca de 80% dos *sites* (Goldfeder et al., 2018). Alguns rastreadores, como a Google e o Facebook, são conhecidos por vincular os perfis de rastreio dos utilizadores a identidades reveladas diretamente pelos utilizadores, mas a maioria dos rastreadores não têm uma relação direta com os utilizadores. No entanto, mesmo esses rastreadores adquirem **IPI**, muitas vezes acidentalmente (Goldfeder et al., 2018).

Estes rastreadores permitem veicular anúncios aos utilizadores para itens que eles sabidamente demonstram interesse em comprar. É uma das formas mais valiosas de publicidade. Quanto mais se puder rastrear um fluxo de pagamento de um utilizador (página do carrinho, página de *checkout*, etc.) maior será o interesse assinalado (Goldfeder et al., 2018). Estes rastreadores são conhecidos como *cookies*, os quais vieram alterar todo o paradigma da navegação na *internet*, para otimizar a experiência do utilizador durante a sua navegação e pesquisa.

A Lei 46/2012, relativa ao tratamento de dados pessoais e à proteção da privacidade no setor das comunicações eletrónicas, entrou em vigor a 29 de agosto de 2012. Esta lei avisa que em território europeu, qualquer *site* tem de avisar e obter consentimento dos utilizadores para poder utilizar *cookies*. No seu artigo 5º podemos ler que "O

armazenamento de informações e a possibilidade de acesso à informação armazenada no equipamento terminal de um assinante ou utilizador apenas são permitidos se estes tiverem dado o seu consentimento prévio, com base em informações claras e completas nos termos da Lei de Proteção de Dados Pessoais, nomeadamente quanto aos objetivos do processamento. (...)” (Artigo 5º da Lei 46/2012). Isto indica que todos os *cookies* de um *site* que não sejam estritamente necessários para a prestação de um serviço pedido pelo utilizador só podem ser guardados no computador do utilizador com a sua autorização prévia e explícita.

Ao longo da próxima secção serão explicados o conceito de *cookie* e os seus vários tipos, bem como os seus objetivos.

2.6.1 Cookies

Cookies são pequenos arquivos de texto que os *sites* colocam nos computadores dos visitantes para armazenar uma série de informação, geralmente específicas para aquele visitante, ou melhor, o dispositivo que eles estão a utilizar para visualizar o *site* (navegador ou *smartphone*). Os *cookies* foram introduzidos em 1994 no navegador Netscape Navigator com o propósito de conveniência do utilizador, nomeadamente lembrando o conteúdo de carrinhos de compras na *web* (Pierson & Heyman, 2011). Eles foram criados para superar uma limitação na tecnologia *web*. As páginas *web* são *stateless*, o que significa que elas não têm memória e não podem facilmente passar informações entre si. Portanto, os *cookies* fornecem uma espécie de memória para as páginas *web*. Os *cookies* permitem ao utilizador fazer *login* numa página, depois mover-se para outras páginas e permanecer *logado*. Permitem ainda que se definam preferências para a exibição de uma página, as quais serão lembradas na próxima vez que se retornar a elas (Cookiepedia, 2019).

Os *cookies* também podem ser usados para ver as páginas que o utilizador visita entre *sites*, o que permite aos anunciantes construir uma imagem dos seus interesses. Depois, quando se chega a um *site* que mostra um dos seus anúncios - eles podem adaptá-lo a esses interesses. Isto é conhecido como "publicidade comportamental".

Quase todos os *sites* utilizam *cookies* de uma forma ou de outra, e cada página que se visita nesses *sites* insere *cookies* no nosso computador e recebe-os de volta a partir dele.

2.6.1.1 Tipos de *Cookies*

Existem dois tipos de *cookies*, os *First Party* e os *Third Party*.

Os ***First Party Cookies*** são *cookies* colocados pelo domínio mostrado na barra de endereço no navegador, e são normalmente utilizados em aplicações de comércio eletrônico, permitindo, por exemplo, a persistência do carrinho de compras (Cahn, et al., 2016). Os ***Third Party Cookies*** são *cookies* colocados por um domínio que é diferente do que é mostrado na barra de endereço do navegador (Cahn, et al., 2016). São rotineiramente implantados por anunciantes *online* e aplicações de rastreamento, como por exemplo a Google Analytics (Cahn, et al., 2016) e usados para monitorar e rastrear o comportamento *online*, pesquisas e *sites* visitados. São colocados em milhares de *sites* que pertencem à rede de publicidade com a finalidade de exibir publicidade "relevante" (Laudon & Traver (2019).

Estes dois tipos de *cookies* dividem-se em duas subcategorias:

- ***Cookies de Sessão*** que permanecem num navegador e retêm as informações do utilizador até que ele seja fechado. Quando uma nova janela do navegador é aberta, o mesmo utilizador é tratado como um novo visitante e deve introduzir as suas credenciais de *login* (Cookiepedia, 2019)
- ***Cookies Persistentes*** que têm uma vida útil designada e permanecem num navegador até que o período transcorra ou o *cookie* seja apagado manualmente. Os *sites* que utilizam *cookies* persistentes vão-se lembrar dos utilizadores mesmo depois de fecharem um navegador. Estes *cookies* permitem funcionalidades como carrinhos de compras persistentes, que retêm produtos adicionados ao carrinho de compras entre sessões (Cookiepedia, 2019).

2.6.1.2 Objetivos dos *cookies*

Uma das exigências das leis sobre os *cookies* é que os utilizadores devem receber informações suficientes sobre o uso de *cookies* para poderem dar ou reter o seu consentimento. Embora não existam requisitos específicos sobre como esta informação deve ser dada, existe um acordo geral de que diferentes tipos de *cookies* podem ser classificados em grupos, de acordo com a sua finalidade. A Cookiepedia (2019)

categoriza, desta forma, os *cookies* de acordo com quatro objetivos, como se apresenta na Tabela 7:

Tabela 7- Objetivos dos *Cookies*

<i>Cookies</i> Estritamente Necessários	Essenciais para que se possa mover pelo <i>site</i> e utilizar as suas funcionalidades. Sem os serviços destes <i>cookies</i> , funcionalidades como o carrinho de compras ou a faturação eletrónica não podem ser fornecidas. Geralmente, são <i>cookies</i> de sessão que expiram ao fechar o navegador, mas nem sempre.
<i>Cookies</i> de Análise/ Performance	Estes <i>cookies</i> recolhem informação sobre como os visitantes utilizam um <i>site</i> , por exemplo, a que páginas os visitantes vão com mais frequência e se recebem mensagens de erro de páginas da <i>web</i> . Estes <i>cookies</i> não recolhem informação que identifique um visitante. Toda a informação que estes <i>cookies</i> recolhem é agregada e, portanto, anónima. É apenas usada para melhorar o funcionamento de um <i>site</i> .
<i>Cookies</i> de Funcionalidade	Estes <i>cookies</i> permitem que o <i>site</i> se lembre das escolhas que o utilizador faz (como o seu nome de utilizador, o idioma ou a região em que está) e fornecem recursos aprimorados e mais pessoais. Por exemplo, um <i>site</i> pode ser capaz de lhe fornecer relatórios meteorológicos locais ou notícias de trânsito, armazenando num cookie a região em que se está atualmente localizado. As informações que estes <i>cookies</i> recolhem podem ser anónimas e não podem acompanhar a atividade de navegação noutros <i>sites</i> .
<i>Cookies</i> de Publicidade	Estes <i>cookies</i> são utilizados para fornecer anúncios mais relevantes para o utilizador e para os seus interesses. São também utilizados para limitar o número de vezes que se vê um anúncio, bem como para ajudar a medir a eficácia de uma campanha publicitária, são normalmente colocados por redes de publicidade com a permissão do operador do <i>site</i> .

Adaptado de Cookiepedia, 2019

Resumindo, os *Cookies*, como anteriormente foi explicado, são uma forma de um *site*, e as pessoas que possuem esse *site*, armazenarem e recuperarem dados sobre o utilizador

ou a sua interação com o *site*. Fazem isso para alterar o que essa pessoa vê, ou registrar a sua atividade para, por exemplo, as páginas que visitam ou quanto tempo passaram num *site* (Pierson; Heyman, 2011). Os *cookies* são fundamentais para a experiência da *web* moderna. Portanto, embora não sejam intrinsecamente maus, há mau uso deles quando levantam preocupações relacionadas com a privacidade dos utilizadores.

Neste capítulo é apresentada a metodologia utilizada nesta dissertação, que é o Estudo de Caso. É apresentada uma contextualização do Estudo de Caso, explanando os métodos utilizados na pesquisa e os instrumentos utilizados para a recolha de dados.

3.1 O Estudo de Caso

O Estudo de Caso constitui-se como um estudo intensivo e detalhado de uma entidade bem definida, um caso único, específico, diferente e complexo (Sousa & Baptista, 2011), tratando-se de uma metodologia aplicada para descrever ou avaliar situações em que o elemento humano está presente (Martins, 2008). Segundo Yin (2001) o Estudo de Caso é um método de investigação que é utilizado com muita frequência, principalmente quando a investigação tem como objetivo conhecer o “como?” e o “porquê?”. Tratando-se de uma investigação centrada em casos especiais, as generalizações neste tipo de investigação devem ser feitas com muita cautela e reserva. Martins (2008), considera que o Estudo de Caso é uma das metodologias mais aplicadas na investigação científica em ciências sociais, tendo sido o seu pioneiro, Yin, em 1994.

Para Yin (2018), existem três condições que definem qual o método que se deve usar numa investigação utilizando o Estudo de Caso:

- a forma como a questão de investigação é colocada;
- o controlo que um investigador tem sobre os eventos em investigação;
- o grau de foco nos eventos contemporâneos, em oposição a eventos inteiramente históricos.

Yin (2018) refere-se ainda a três princípios fundamentais para a realização de um Estudo de Caso:

- **1º princípio:** recurso a várias fontes de dados. Para recorrer a várias fontes de dados, é recorrente abordar a triangulação de dados. A triangulação de dados é a determinação da convergência dos dados recolhidos a partir de diferentes fontes de evidência, para avaliar a força de um Estudo de Caso e também para aumentar a validade das medidas utilizadas no Estudo de Caso (Yin, 2018).
- **2º princípio:** criação de uma base de dados. Num Estudo de Caso, as bases de dados podem ser compostas por diferentes formatos de registos, podendo ser documentos, tabelas, relatos, em papel ou em formato digital.

- **3º princípio:** cadeia de evidências. Este princípio forma o investigador para manter uma cadeia de evidências, cujo objetivo é auxiliar o leitor, apresentando de forma coerente que evidências levaram às conclusões.

Os estudos de caso podem ser classificados pelo número de casos, podendo ser um único caso ou um caso múltiplo:

“A escolha entre projetos de caso único ou de casos múltiplos permanece dentro da mesma estrutura metodológica - e nenhuma distinção muito ampla é feita entre o assim chamado estudo de caso clássico (isto é, único) e os estudos de casos múltiplos. A escolha é considerada uma escolha do projeto de investigação, com as duas sendo incluídas no âmbito da estrutura do estudo de caso.” (Traduzido de Yin, 2001, p.68)

Os Estudos de Caso podem ser exploratórios, descritivos ou explanatórios (Yin, 2001). O objetivo dos estudos exploratórios é definir as questões ou as hipóteses para uma investigação posterior. Os estudos descritivos representam uma descrição o mais completa possível de um fenómeno inserido no seu contexto. Por fim, os estudos explanatórios procuram informações que possibilitem o estabelecimento de relações de causa-efeito (Yin, 2001). Podemos afirmar ainda que o Estudo de Caso exploratório é caracterizado pela investigação de algum objeto de estudo que possui poucas informações. O Estudo de Caso descritivo vai de encontro às análises de dados quantitativas e qualitativas e procede ao levantamento dos resultados obtidos e ao porquê desses dados. O Estudo de Caso explanatório tem como finalidade informar e explicar fenómenos (Yin, 2018).

Para Dalfovo (2008), a abordagem quantitativa refere-se a tudo aquilo que é mensurável. Esta abordagem apela ao uso da quantificação, tanto na recolha dos dados, como no tratamento, utilizando técnicas estatísticas que objetivam resultados, evitando possíveis distorções na análise e interpretação dos dados, fornecendo assim uma maior margem de segurança (Diehl, 2004). A recolha de dados nesta abordagem costuma ser realizada através de questionários que, depois de analisados os seus dados, os resultados são apresentados através de tabelas e gráficos (Dalfovo, 2008). Já a abordagem qualitativa, não podendo ser traduzida em números, pretende verificar uma relação entre a realidade e o objeto de estudo, podendo ser obtidas várias interpretações de uma análise indutiva

por parte do investigador (Dalfovo, 2008). A investigação qualitativa “trabalha com valores, crenças, hábitos, atitudes, representações, opiniões e adequa-se a aprofundar a complexidade de fatos⁵ e processos particulares e específicos a indivíduos e grupos.” (Prota *et al.*, 1999, p.135)

Segundo Gil (1999, p.128), o questionário é considerado “como a técnica de investigação composta por um número mais ou menos elevado de questões apresentadas por escrito às pessoas, tendo por objetivo o conhecimento de opiniões, crenças, sentimentos, interesses, expectativas, situações vivenciadas.” Num questionário, as questões colocadas devem ser diretas e focadas no tema que está a ser estudado, portanto, é necessário ter objetivos bem definidos e presentes nas questões. Hill & Hill (2009), afirmam que o questionário consiste em colocar, a um conjunto de pessoas, um conjunto de questões relativas à sua situação profissional e social, às suas opiniões ou atitudes e ao seu nível de conhecimento ou de consciência de um acontecimento ou problema. Ou seja, em resumo, o questionário é o instrumento de investigação indicado quando o objetivo é recolher informações quanto aos comportamentos, valores ou opiniões de um conjunto de inquiridos (Hill & Hill, 2009). Gil (1999) aponta algumas vantagens na utilização do questionário, como a garantia de anonimato, a possibilidade de o inquirido responder quando lhe é conveniente, poucos gastos com pessoal, atinge um grande número de pessoas e abrange uma área geográfica mais ampla. Moysés & Moori (2007) acrescentam ainda a possibilidade de resposta a partir de diversas tecnologias (*smartphones*, computadores fixos, portáteis, *tablets*...), a ausência de deslocamentos e a economia de tempo.

3.1.1 Caso Presente

No estudo apresentado nesta dissertação, a estratégia de pesquisa, refere-se ao “estudo de caso”, com tipologia de “caso múltiplo” de características descritivas, tendo sido utilizada uma abordagem quantitativo-qualitativa. Este estudo classifica-se como descritivo uma vez que o objetivo principal é perceber a utilização dos Big Data no contexto do comércio eletrónico, quer do ponto de vista dos utilizadores, quer do ponto de vista das organizações. Foram recolhidos dados a partir de um questionário, que pretende avaliar o nível de consciencialização do utilizador *online* sobre a recolha dos seus dados pessoais em contexto de comércio eletrónico. Dada a necessidade de inquirir um número considerável de pessoas, o questionário, no caso deste estudo, é o método científico que

⁵ Do Português do Brasil. Em Português de Portugal, “facto”.

mais se aplica, visto que este instrumento permite, de acordo com (Hill & Hill, 2009), a recolha de dados de forma rápida e padronizada, facilitando assim a análise dos dados. Foi ainda efetuada a análise das políticas de privacidade e *cookies* de um conjunto selecionado de *sites* de comércio eletrónico, que permitirá concluir quais os dados pessoais dos utilizadores *online* recolhidos por parte das organizações e qual o seu objetivo.

3.2 Abordagem à investigação e métodos de recolha de dados

No que concerne à abordagem utilizada neste estudo, foi primeiramente feita uma revisão narrativa de literatura que se focou em vários objetivos específicos, nomeadamente:

- perceber o conceito de Big Data;
- identificar as suas características, principais tecnologias e benefícios ou desafios;
- definir quais as questões legais e éticas da utilização dos dados do utilizador online em ambiente de comércio eletrónico.

Neste procedimento, que foi trabalhado no Capítulo 2 desta dissertação, coligimos diferentes autores que serviram como suporte teórico para que, num segundo passo, se conseguisse aplicar o conhecimento adquirido na elaboração do questionário aos utilizadores *online*, na análise às políticas de privacidade anunciadas, e aos *cookies* utilizados por *sites* de comércio eletrónico.

Para este estudo, foram recolhidos dados a partir da aplicação de um questionário *online*, referente à consciencialização dos utilizadores *online* quanto à recolha dos seus dados pessoais por parte de *sites* de comércio eletrónico, bem como uma análise dos *cookies* utilizados num conjunto selecionado de *sites* de comércio eletrónico e às suas políticas de privacidade, de forma a concluir quais os *cookies* e tipos de informação que esses mesmos *sites* recolhem dos utilizadores. Estes dois instrumentos de recolha de dados são descritos nas seções seguintes e os dados recolhidos serão analisados no capítulo seguinte.

3.2.1 Questionário *online*

No caso deste estudo, a primeira recolha de dados foi realizada por meio da aplicação de um questionário *online* implementado no *Google Forms*. O questionário teve como objetivo geral avaliar o nível de consciencialização dos utilizadores *online* quanto à partilha e recolha dos seus dados por parte de *sites* de comércio eletrónico. O questionário, que se encontra no Apêndice I, é composto por 6 secções e apresenta 14

(catorze) questões, sendo que 13 (treze) são de resposta fechada, incluindo uma questão feita com uma escala de Likert de 5 pontos, e 1 (uma) de resposta aberta (última questão, referente a algum comentário adicional).

A primeira secção inclui 4 questões e tem como objetivo determinar as características da amostra. Para essa caracterização, foram definidas as variáveis *Idade*, *Sexo*, *Grau de Escolaridade* e *Tipo de Dispositivo* que mais utiliza para aceder à Internet.

A segunda secção, composta por três questões, pretende definir o nível de conhecimento do utilizador *online* relativamente ao conceito de dados pessoais. Na primeira questão desta secção, foi pedido aos inquiridos que selecionassem entre as opções *Nome*, *Morada*, *Número de Telemóvel*, *Email*, *Dados de Contas Bancárias*, *Género*, *Idade*, *Ocupação*, *Localização*, *Gostos (likes)*, *Impressões Digitais*, *Nível de Educação*, *Transações Financeiras*, *Religião*, aquelas que considerassem serem dados pessoais. Poderiam ser selecionadas todas as opções desejadas pelos inquiridos. Na segunda questão desta secção foi questionado se os inquiridos costumam fornecer os seus dados pessoais em contexto de comércio eletrónico. De seguida, aos inquiridos que respondessem que costumavam fornecer os seus dados pessoais em contexto de comércio eletrónico, pedia-se que selecionassem as 3 principais razões que os levavam a fornecer os seus dados pessoais, de entre as seguintes opções: Fazer um pagamento; Obter um serviço; Poupar tempo na próxima visita; Obter um serviço gratuito; Conseguir a entrega de uma compra.

Na terceira secção pretende-se avaliar o comportamento do utilizador perante a partilha dos seus dados pessoais em ambiente de comércio eletrónico. Nesta secção do questionário foram feitas duas questões, sendo a primeira “Que controlo os inquiridos sentem que têm sobre a informação que disponibilizam *online*” Foi também pedido aos inquiridos, que mostrassem se concordavam ou não com um conjunto de afirmações relacionadas com a forma como controlam a monitorização das suas atividades online. A quarta secção tem como objetivo determinar o nível de conhecimento do utilizador *online* quanto aos seus direitos inerentes no RGPD. Nesta secção questionou-se se os inquiridos tinham conhecimento sobre o Regulamento Geral de Proteção de Dados e em caso afirmativo, teriam de assinalar quais de três afirmações eram corretas e quais eram incorretas.

Quanto à quinta secção, pretende-se definir o nível de concordância dos inquiridos perante as opções que os *sites* de comércio eletrónico oferecem para a sua utilização.

Nesta secção do questionário pediu-se aos inquiridos que respondessem se concordavam ou não com um conjunto de afirmações sobre a monitorização que os *sites* de comércio eletrónico fazem sobre as suas atividades online. Com a questão seguinte do questionário, utilizando uma escala de Likert de 5 pontos (1- Discordo Totalmente, 2- Discordo Parcialmente, 3- Indiferente, 4- Concordo Parcialmente, 5- Concordo Totalmente), pede-se aos inquiridos que avaliem o seu sentimento em relação ao fornecimento e recolha constante de dados pessoais em contexto de comércio eletrónico. A última questão do questionário exigia uma resposta aberta facultativa por parte dos inquiridos para introdução de comentários adicionais.

De forma a tornar mais claros os objetivos deste questionário, foi elaborada a Tabela 8, que apresenta uma correspondência entre os objetivos definidos para o questionário e as questões nele incluídas.

Tabela 8- Correspondência: Objetivos e Questões do Questionário

Objetivo Geral	Secções do Questionário	Objetivos Específicos	Nº das questões
Avaliar o nível de consciencialização do utilizador em relação à partilha e recolha dos seus dados em ambiente de comércio eletrónico	Caracterização da amostra	Apresentar características gerais da amostra por meio das respostas recebidas	1, 2, 3, 4
	Nível de Conhecimento do Utilizador <i>Online</i> em relação ao conceito de dados pessoais	Apontar os dados que os utilizadores consideram ser pessoais e qual o motivo para a sua partilha para com os <i>sites</i> de comércio eletrónico	5, 6, 7
	Avaliação do comportamento do utilizador perante a partilha dos seus dados	Perceber quais os comportamentos dos utilizadores perante a sua noção de controlo sobre os seus dados	8, 9
	Nível de conhecimento perante os seus direitos inerentes no RGDP	Entender o nível do conhecimento do utilizador <i>online</i> em relação ao Regulamento Geral de Proteção de Dados	10, 11
	Nível de concordância perante as opções que os <i>sites</i> oferecem para a sua utilização	Compreender qual a posição dos utilizadores perante a recolha e utilização dos seus dados pessoais	12, 13
	Aspectos adicionais		14

Este questionário foi divulgado para os utilizadores *online* durante 34 dias, entre 24 de março de 2020 e 27 de abril de 2020. O questionário foi divulgado via *e-mail* e através das redes sociais (Facebook, Instagram e WhatsApp). Neste questionário, a amostra foi obtida através de um método não-probabilístico e foi utilizado o tipo de amostragem conhecido como *Bola de Neve*.

3.2.2 Cookies utilizados por sites de comércio eletrónico

Quanto à segunda recolha de dados, foi feito um estudo relativo às políticas de privacidade e às *cookies* utilizadas pelos *sites* de comércio eletrónico, a partir de uma análise a um grupo selecionado de *sites*.

Os objetivos deste segundo estudo focaram-se em 6 pontos:

1. Selecionar um conjunto de *sites* do âmbito de comércio eletrónico;
2. Analisar os *cookies* presentes nesses *sites*;
3. Organizar os resultados por tipo de *cookie* e objetivo;
4. Comparar quais os *cookies* mais usados;
5. Examinar a sua função;
6. Concluir quais os dados mais retirados nesse conjunto de *sites*.
7. Perceber, segundo a Política de Privacidade de cada *site*, se as informações fornecidas correspondem com a função dos *cookies* encontrados no *site* a partir deste estudo

Como critério para a seleção dos *sites*, optou-se por considerar os *sites* mais utilizados em Portugal. Além disso, considerou-se importante analisar *sites* com diversos modelos de negócio de comércio eletrónico, tendo sido selecionados os seguintes: retalhista, C2C (Consumer to Consumer), rede social e portal. Foi necessário identificar uma plataforma que apresentasse os *sites* de comércio eletrónico mais utilizados em Portugal. Desta forma, recorremos ao SimilarWeb⁶. Nesta plataforma, foi possível aceder a uma semana gratuita de utilização, tendo sido possível aceder a uma lista ordenada dos *sites* mais utilizados em Portugal. Nesta lista foram selecionados aqueles que se inseriam na categoria de comércio eletrónico e foram tidos em conta os respetivos modelos de negócio de comércio eletrónico já referidos acima.

⁶ www.similarweb.com - empresa que disponibiliza uma plataforma que utiliza tecnologias de Big Data para recolher dados, medir, analisar e fornecer estatísticas do envolvimento dos utilizadores em *sites* e aplicações móveis.

De seguida utilizou-se a ferramenta de pesquisa disponibilizada pela Cookiepedia⁷, visto que esta plataforma permite aceder aos dados secundários relativos aos *cookies* utilizados por cada *site*. A Cookiepedia é a maior base de dados de pré-categorias de *cookies* e tecnologias de rastreio *online*. O *site* Cookiepedia foi configurado para preencher uma lacuna de informações sobre o que os *cookies* fazem, quem os utiliza e para que fins, como gerir *cookies* e é mantido pela OneTrust, uma empresa de software de gestão de privacidade.

Numa primeira fase, foram introduzidos na ferramenta de pesquisa da Cookiepedia os *sites* de comércio eletrónico presentes na lista dos *sites* mais utilizados em Portugal, fornecida pelo SimilarWeb, para ver se era possível obter dados. Estes testes foram necessários visto que alguns *sites* possuem bloqueadores nos seus *cookies*, não sendo possível diretamente aceder à sua lista de *cookies*. Na Figura 3 apresenta-se a ferramenta de pesquisa fornecida pela Cookiepedia, onde é possível escolher se se pretende analisar um *site*, um *cookie* ou um *3rd party host*.

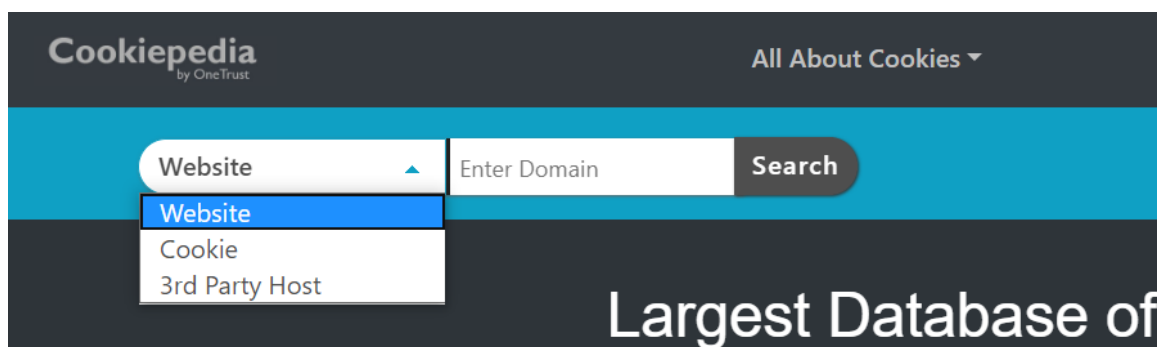


Figura 3- Ferramenta de pesquisa fornecida pela Cookiepedia

Apresenta-se, na Figura 4, um exemplo de um *site* de comércio eletrónico para o qual a ferramenta de pesquisa da Cookiepedia não consegue recolher nenhum dado.



Figura 4- Exemplo de *site* em que a Cookiepedia não consegue obter dados

⁷ <https://cookiepedia.co.uk/>

Depois de fazer estes testes com vários *sites* de comércio eletrónico presentes na lista dos *sites* mais utilizados em Portugal obteve-se uma amostra de conveniência de 12 *sites* de comércio eletrónico, que são possíveis de analisar através da ferramenta de pesquisa da Cookiepedia, os quais são apresentados na Tabela 9.

Tabela 9- *Sites* seleccionados no estudo de *Cookies*

B2C (retalhista)	C2C	Redes Sociais	Portal
Amazon	Custo Justo	Facebook	Google
Continente	Ebay	LinkedIn	
Decathlon			
Fnac			
SportZone			
Wook			
Worten			

Apresenta-se de seguida uma breve descrição de cada um dos *sites*:

- A amazon.com⁸ é uma empresa de tecnologia sediada nos Estados Unidos que se foca em comércio eletrónico, computação em nuvem, *streaming* digital e inteligência artificial.
- O Continente⁹ foi a primeira cadeia de hipermercados em Portugal, e mantém-se ainda hoje como uma referência no sector de retalho alimentar do país.
- A Decathlon¹⁰ é um dos maiores retalhistas de artigos desportivos do mundo.
- A Fnac¹¹ é uma cadeia de lojas que oferece produtos culturais e tecnológicos.
- A SportZone¹² é uma cadeia de lojas de vestuário e produtos desportivos portuguesa do grupo Sonae.
- A Wook¹³ é uma livraria portuguesa *online*.
- A Worten¹⁴ é uma empresa portuguesa, pertencente ao grupo Sonae, de eletrónica de consumo e do entretenimento.

⁸ <https://www.amazon.com/>

⁹ <https://www.continente.pt/pt-pt/public/Pages/homepage.aspx>

¹⁰ <https://www.decathlon.com/>

¹¹ <https://www.fnac.pt/>

¹² <https://www.sprintersports.com/pt/>

¹³ <https://www.wook.pt/>

¹⁴ <https://www.worten.pt/>

- O CustoJusto¹⁵ é um *site* de Classificados onde se encontra todo o tipo de produtos à venda, desde anúncios de casas, carros e motos usadas, ofertas de emprego, móveis e eletrodomésticos para casa.
- O eBay¹⁶ é uma empresa de comércio eletrónico fundada nos Estados Unidos, ligado à venda e compra de bens.
- O Facebook¹⁷ é a maior rede social, fundada por Mark Zuckerberg, proprietário de outras redes sociais como o Instagram, Snapchat e WhatsApp.
- O LinkedIn¹⁸ é uma rede social direcionada a negócios.
- A Google¹⁹ é uma empresa multinacional de serviços *online* e software dos Estados Unidos que hospeda e desenvolve uma série de serviços e produtos baseados na *internet* e gera lucro principalmente através da publicidade pelo AdWords.

Com os dados fornecidos pela ferramenta de pesquisa da Cookiepedia, foi criada uma base de dados em Excel, na qual foram organizados os resultados dos *cookies* por tipo e objetivo, tal como apresentada na secção 2.6.1. Com base nesta base de dados criada em Excel, foi também possível analisar os *cookies* mais utilizados e mais importantes nestes *sites*, sendo assim possível dar exemplos da função exata de alguns *cookies*.

É de salientar que o RGPD obriga os *sites* a pedir a autorização do uso de *cookies* aos utilizadores online. Se se desativar ou não autorizar os *cookies*, a navegação no *site* poderá não ser otimizada e algumas das funcionalidades disponíveis no *site* podem não funcionar corretamente. Foi assim feita uma análise sobre a Política de Privacidade apresentadas por todos os *sites*, de modo a perceber qual a utilização que as empresas afirmam fazer com os dados recolhidos, nomeadamente através dos *cookies*. Para cada *site* foram analisados os seguintes aspetos:

- O que acontece ao entrar no *site*
- Quais os dados ou tipos de dados que os *sites* informam recolher
- Informação específica sobre *cookies*

¹⁵ <https://www.custojusto.pt/>

¹⁶ <https://www.ebay.com/>

¹⁷ <https://www.facebook.com/>

¹⁸ <https://www.linkedin.com>

¹⁹ <https://www.google.com/>

- Como informam os *sites* que utilizam esses dados que recolhem
- Informação fornecida ao utilizador sobre como controlar os dados fornecidos

Por fim, pretendia-se analisar alguns dos *cookies* presentes nos *sites* analisados. Alguns dos *sites* analisados forneciam, na sua página de política de privacidade, listas com os *cookies* utilizados nos seus *sites*. Naqueles que não o faziam, foi utilizada a base de dados feita em Excel a partir dos dados recolhidos pela ferramenta de pesquisa da Cookiepedia. Ao analisar estes *cookies*, pretendia-se perceber a função exata de um cookie em específico num *site*. Desta forma, na ferramenta de pesquisa da Cookiepedia, optava-se pela opção “Cookie” e inseria-se o nome do cookie na barra de pesquisa. Apresenta-se na Figura 5 um exemplo.

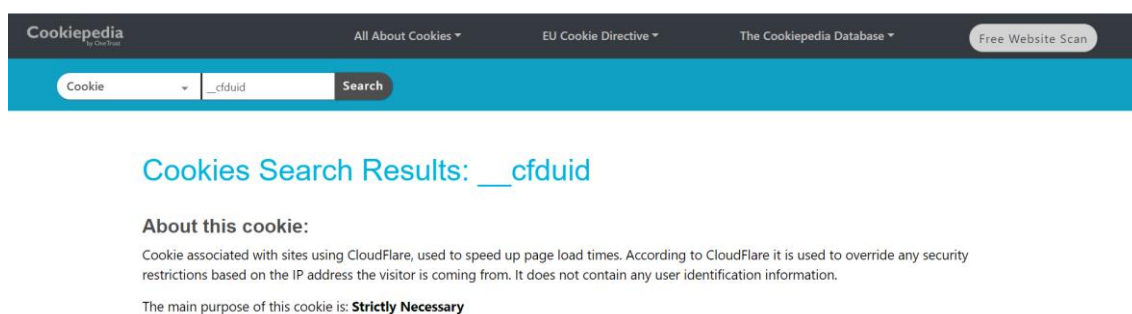


Figura 5- Exemplo de pesquisa na Cookiepedia de um *cookie* específico

Por norma, quando a ferramenta de pesquisa da Cookiepedia não reconhece um *cookie*, classifica-o como desconhecido e isso quer dizer que este é unicamente utilizado pela empresa dona do *site*, tendo sido a mesma a criá-lo, ou criado por uma empresa particular, para fins específicos.

Assim, tornou-se possível uma comparação entre os resultados obtidos e as informações fornecidas sobre os *sites* relativamente ao tratamento de dados pessoais.

Os resultados destas análises são apresentados na secção 4.2.

Neste capítulo apresentamos os resultados obtidos por meio da análise dos dados recolhidos via os dois instrumentos de recolha de dados explanados no capítulo anterior.

4.1. Análise do questionário *online*

Ao todo, o questionário recebeu 160 respostas, todas válidas. Analisemos seguidamente as respostas obtidas de acordo com cada secção do questionário.

4.1.1. Caracterização da Amostra

Observa-se (Tabela 10) que, dos 160 inquiridos, a maioria se encontra no intervalo de idades 24 a 33 anos (37,5%, 60 inquiridos), com uma pequena diferença para intervalo de idades 14 a 23 anos (30%, 48 inquiridos), seguindo-se o intervalo de idades 44 a 53 anos (23,1%, 37 inquiridos). Os restantes intervalos de idades apresentam uma percentagem de inquiridos inferior a 5%. A média de idades é 38,2 anos e a moda 23 anos.

Tabela 10- Faixa etária dos inquiridos

Faixa etária	Percentagem de Inquiridos
14 - 23 anos	30,0%
24 - 33 anos	37,5%
34 - 43 anos	5,0%
44 - 53 anos	23,1%
54 - 63 anos	3,8%
64 - 73 anos	0,6%

No que concerne ao *Sexo* dos inquiridos, 53% (85 inquiridos) são do *Sexo Masculino* e 47% (75 inquiridos) são do *Sexo Feminino* (Figura 6).

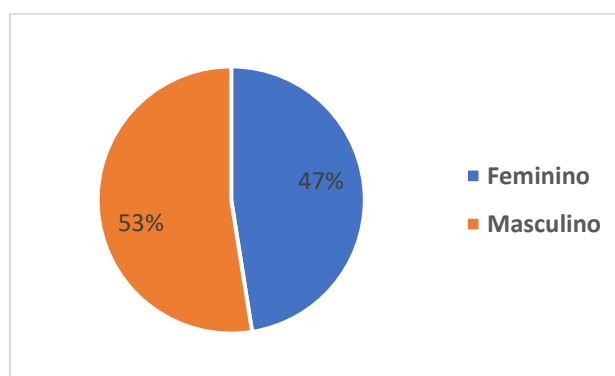


Figura 6- Sexo dos inquiridos

Quanto ao *Grau de Escolaridade* dos inquiridos, podemos observar (Tabela 11) que quase metade dos inquiridos (49%, 78 inquiridos) têm *Licenciatura*. De seguida, temos o *Ensino Secundário*, com 24% (38 inquiridos) e em terceiro lugar *Mestrado*, com 19% (30 inquiridos). Os restantes *Graus de Escolaridade* apresentam percentagens inferiores a 4% de inquiridos.

Tabela 11- Grau de Escolaridade dos Inquiridos

Grau de Escolaridade	Percentagem de Inquiridos
Ensino Básico – 2ºCiclo	4,4% 7 inquiridos
Ensino Secundário	23,8% 38 inquiridos
Curso Profissional	1,3% 2 inquiridos
Licenciatura	49,4% 79 inquiridos
Bacharelato	1,3% 2 inquiridos
Mestrado	19,4% 31 inquiridos
Doutoramento	0,6% 1 inquirido

Por último, foi questionado qual o *Tipo de Dispositivo* que os inquiridos mais utilizam para aceder à *Internet*. Pelos dados obtidos percebemos que o *Smartphone* é o dispositivo que os inquiridos mais utilizam (62%, 100 inquiridos), seguindo-se o *Computador* (35%, 56 inquiridos) e por fim o *Tablet* (3%, 4 inquiridos) (Figura 7).

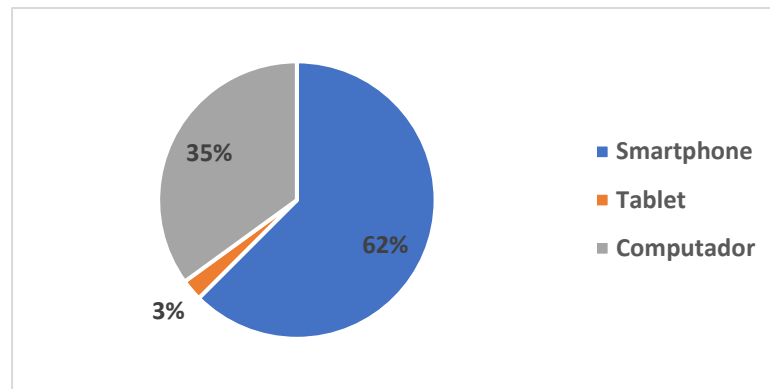


Figura 7- Tipo de Dispositivo que os inquiridos mais utilizam para aceder à Internet

4.1.2. Nível de conhecimento do utilizador online em relação ao conceito de dados pessoais

Apesar de todas as opções apresentadas no questionário se referirem a Informações Pessoais Identificáveis (IPI), verifica-se que os inquiridos não têm consciência desse facto. *Número de Telemóvel* (96,3%, 155 inquiridos), *Morada* (94,4%, 152 inquiridos), *Dados de Contas Bancárias* (92,5%, 149 inquiridos) são os dados que a maior percentagem dos inquiridos considera como sendo pessoais, logo seguidos de *Nome* (86,3 %, 139 inquiridos), *Impressões digitais* (84,5%, 136 inquiridos) e *Email* (84,5%, 136 inquiridos). Menos de 50% dos inquiridos consideram que a *Ocupação* (44,1%, 71 inquiridos), *Género* (42,2%, 68 inquiridos), *Religião* (36%, 58 inquiridos), *Nível de Educação* (36%, 58 inquiridos), *Gostos* (29,2%, 47 inquiridos) e *Tipo de Navegador* (20,5%, 33 inquiridos) são **IPI** (Figura 8).

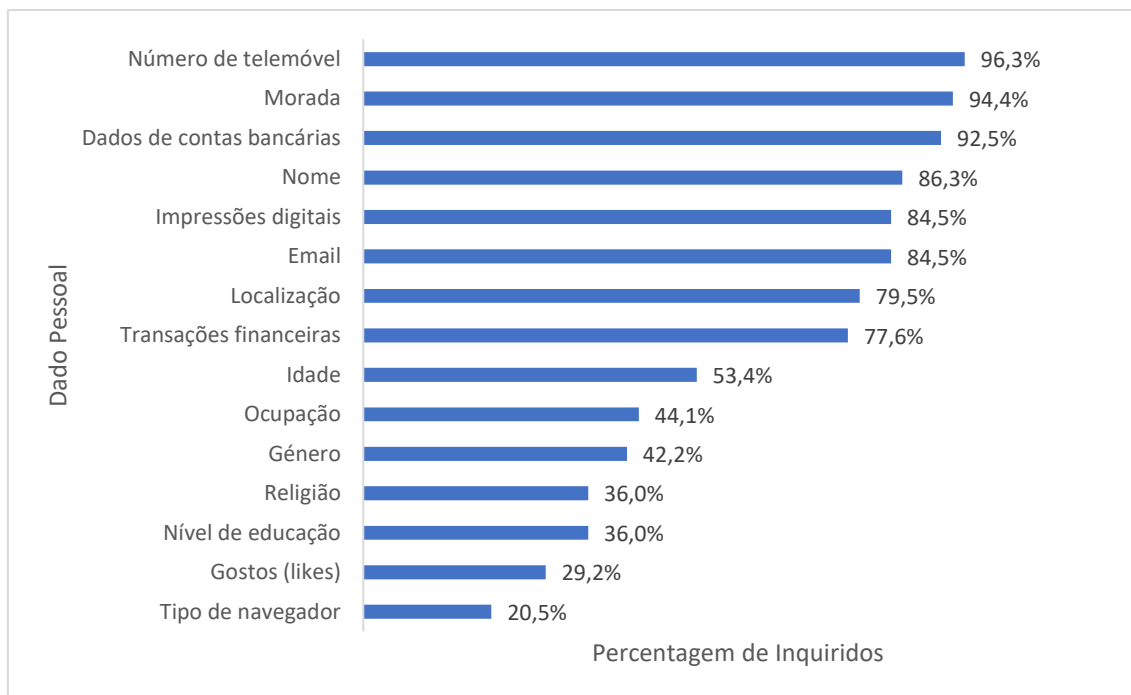


Figura 8- Dados que os inquiridos consideram como pessoais

Quando questionado se costumam fornecer os seus dados pessoais em contexto de comércio eletrónico, 54,4% dos inquiridos (87 inquiridos) responderam afirmativamente, e 45,6% negativamente (73 inquiridos) (Figura 9).

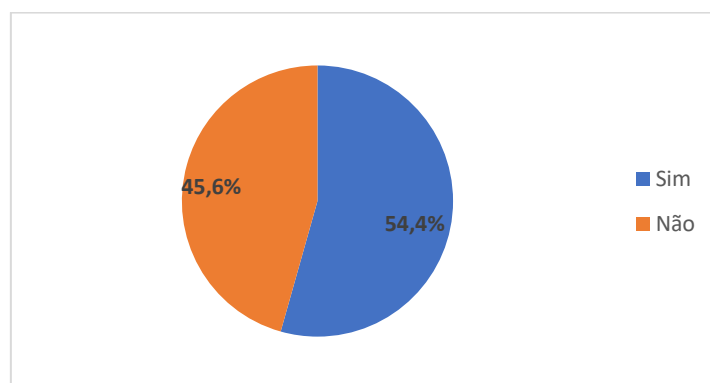


Figura 9- Percentagem de inquiridos que costumam fornecer os seus dados pessoais em contexto de comércio eletrónico

De entre os inquiridos que fornecem os seus dados pessoais, as razões apontadas para esse facto são para “Fazer um pagamento” (96,6%, 84 inquiridos), para “Conseguir a entrega de uma compra” (82,8%, 72 inquiridos) e para “Obter um serviço” (80,5%, 70 inquiridos). Poucos dos inquiridos fornecem os seus dados para “Poupar tempo na próxima visita” (25,3%, 22 inquiridos) ou para “Obter um serviço gratuito” (17,2%, 15 inquiridos) (Figura 10).

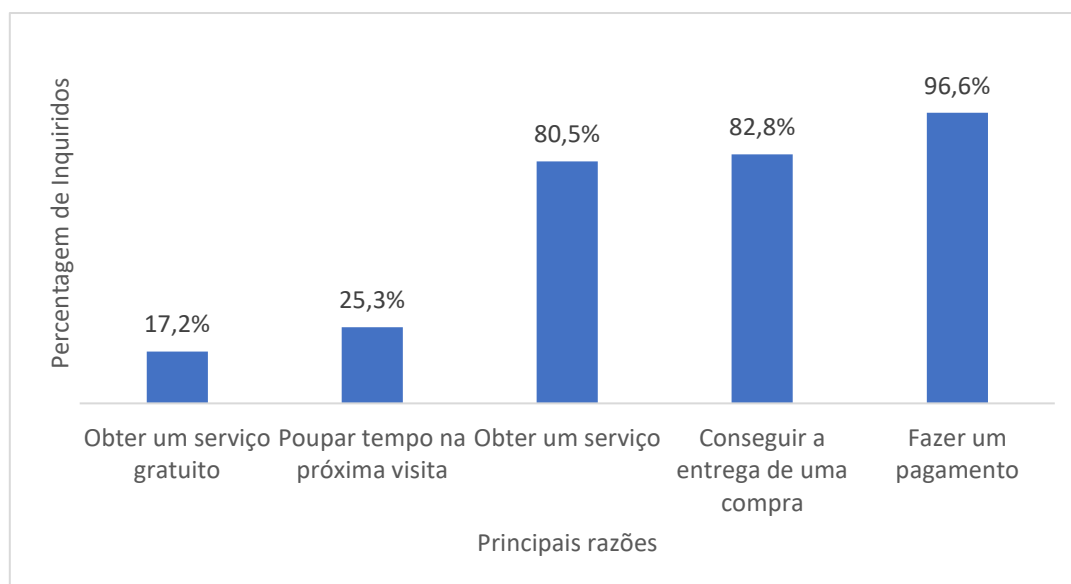


Figura 10- Principais razões para fornecer dados pessoais

4.1.3. Avaliação do comportamento do utilizador perante a partilha dos seus dados

Os resultados obtidos mostram que a maioria dos inquiridos (36,9%, 59 inquiridos) considera ter um controlo *Parcial* sobre a informação que disponibilizam *online*, 33,8% (54) dos inquiridos acham que *Depende do site ou aplicação* e 20,6% (33) inquiridos sentem que não têm *Nenhum* controlo sobre a informação fornecida. São menos de 9% (14 inquiridos) que sentem um controlo *Total* sobre a informação disponibilizada ou *Não sabem* o controlo que têm (Figura 11).

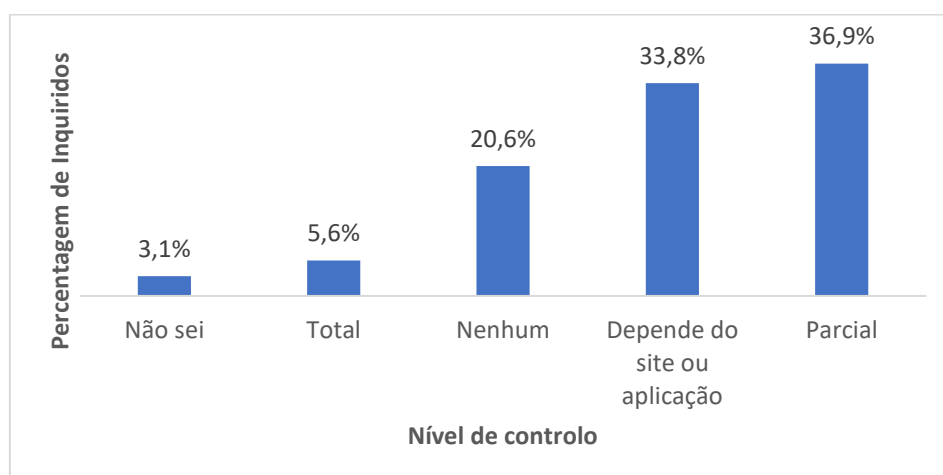


Figura 11- Nível de controlo sobre a informação disponibilizada online

São mais os inquiridos que “têm por hábito aceitar os *cookies* quando acedem a um *site*” (57,5%, 92 inquiridos). Já quanto à questão “Utiliza software que evita que as suas atividades online sejam monitorizadas” os inquiridos quase se encontram igualmente divididos, visto que 51,3% (82) inquiridos afirmaram utilizar algum software e 48,8% (78 inquiridos) não usam. O mesmo acontece quanto à “Mudança das definições de privacidade do navegador que utilizam”, sendo que 52,5% (84) dos inquiridos afirmam já o ter feito, e 47,5% (76 inquiridos) não o fazerem. 77,5% (124) dos inquiridos afirma “Evitam alguns *sites* de comércio eletrónico por receio de que a sua atividade seja monitorizada” (Figura 12).

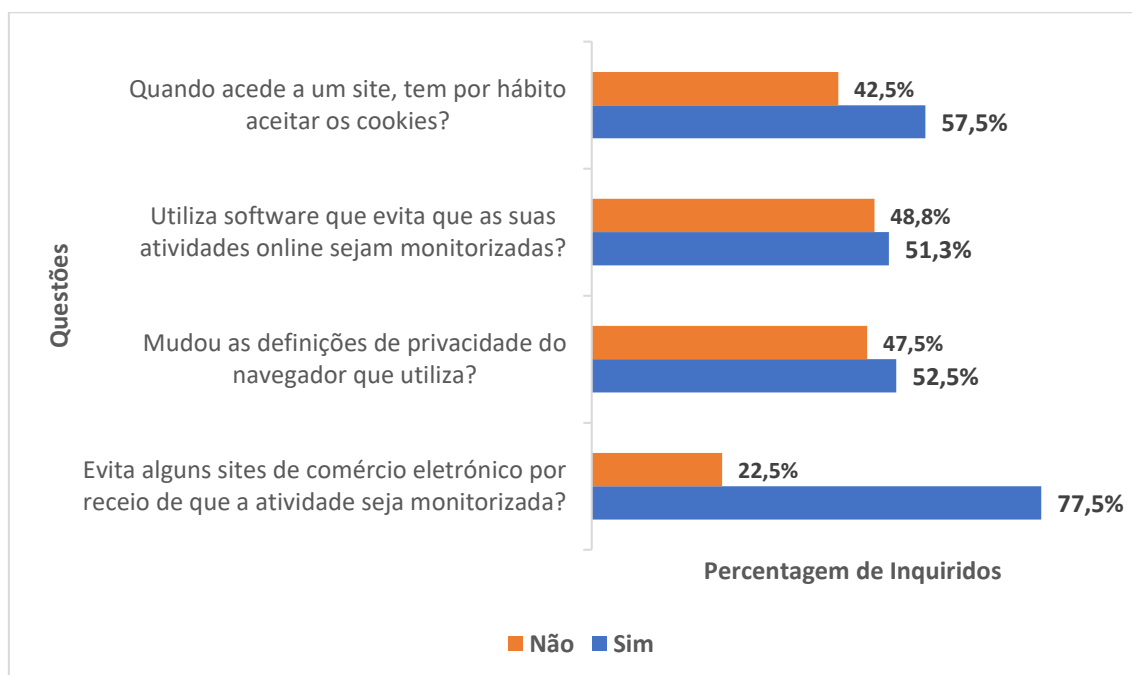


Figura 12- Considerar aceitável opções que os *sites* oferecem para a sua utilização

4.1.4. Nível de conhecimento perante os seus direitos inerentes no RGDP

Ao questionar os inquiridos sobre se tinham conhecimento do que era o Regulamento Geral de Proteção de Dados (RGPD), 21% (34 inquiridos) responderam que não, e 79% (126 inquiridos) responderam que sim (Figura 13)

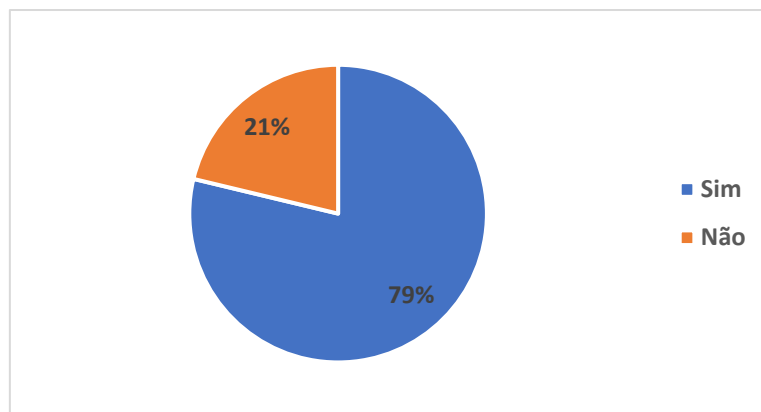


Figura 13- Conhecimento do que é o Regulamento Geral de Proteção de Dados

Dos 126 inquiridos que dizem conhecer o RGPD, (Figura 14), vê-se que apenas 93,7% (118) dos inquiridos sabe que “Os cidadãos podem exigir que os seus dados pessoais sejam eliminados”. 64,3% (81) dos inquiridos não sabe que “O RGPD obriga as empresas a registar os dados pessoais dos cidadãos num formato que permita que sejam enviados para outra empresa, facilitando a sua migração”. 61,1% (77) dos inquiridos não sabem que “O Regulamento se aplica em *sites* de comércio eletrónico fora da União Europeia”. Apesar de todas as afirmações apresentadas serem verdadeiras, com estes resultados podemos concluir que os inquiridos não conhecem devidamente o conteúdo do RGPD.

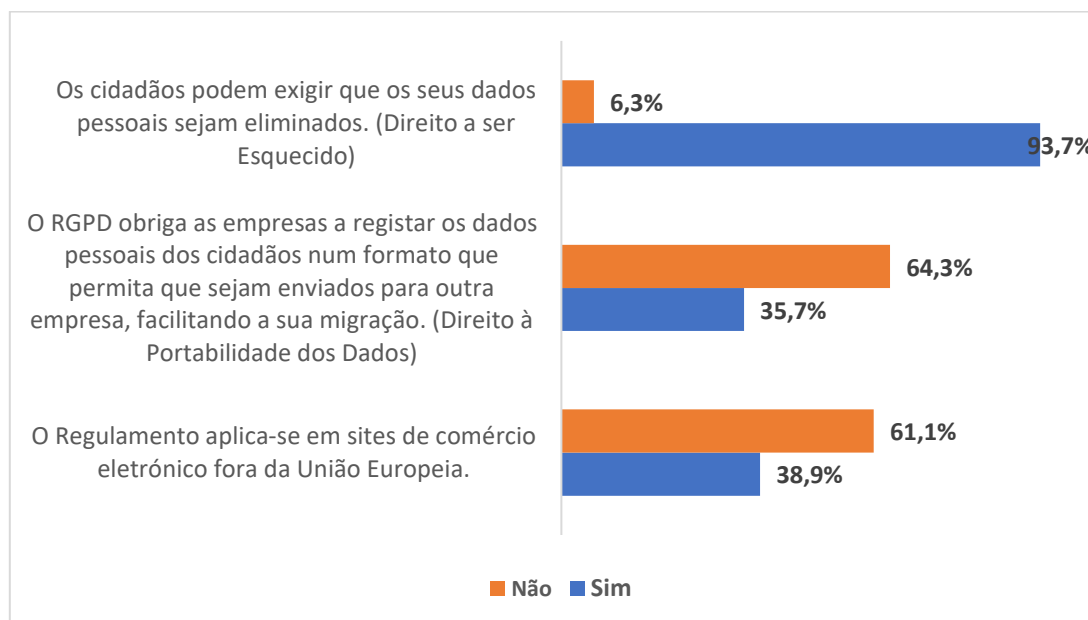


Figura 14- Opção correta para um conjunto de questões sobre o RGPD

4.1.5. Nível de concordância perante as opções que os *sites* oferecem para a sua utilização

Numa primeira visão, percebemos que os inquiridos não concordam com nenhuma das condições que os *sites* de comércio eletrónico oferecem para a sua livre utilização (Figura 15). A maioria dos inquiridos (91,9%, 147 inquiridos) não concorda em ter de “Pagar para não ser monitorizado enquanto navega num *site*”, 97,5% (156) inquiridos não concorda “Que os *sites* de comércio eletrónico partilhem os seus dados como utilizador sem a sua permissão”, 88,1% (141) inquiridos não concorda “Que os *sites* de comércio eletrónico utilizem os seus dados como utilizador para benefício próprio” e 77,5% (124) dos inquiridos não concorda “Que a sua atividade online seja monitorizada em troca de acesso livre a um *site*”.

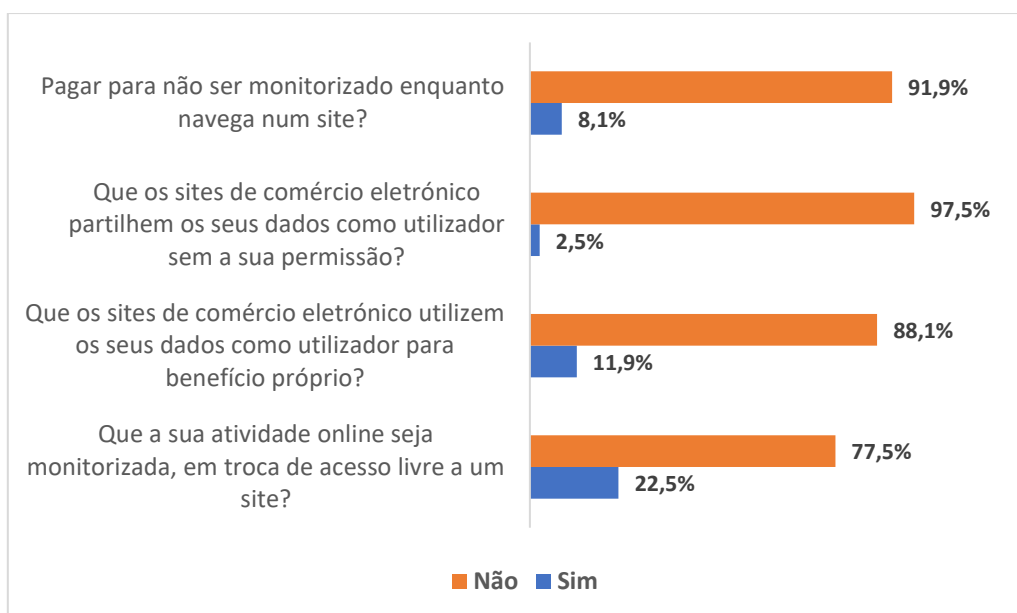


Figura 15- Nível de concordância dos inquiridos sobre as condições que os *sites* oferecem para a sua utilização

Na tabela 12 apresenta-se o resumo das respostas obtidas pelos inquiridos quando questionados sobre a sua opinião acerca de 5 afirmações relacionadas com os seus sentimentos relativamente ao terem de fornecer informações pessoais *online*. Os inquiridos concordam genericamente com todas as afirmações (as médias são superiores a 3). Podemos observar que as afirmações “Sinto que tenho de fornecer cada vez mais informação pessoal” e “Sinto-me invadido ao fornecer informação pessoal”, apresentam um nível de concordância mais claro (médias e medianas mais altas). Os inquiridos apresentam algum nível de concordância (média de 3.33) com a afirmação “Importo-me

de fornecer informação pessoal em troca de serviços online gratuitos, como o email”. De qualquer modo, acabam por reconhecer (médias de 3.34 e 3.35, respetivamente) que “Fornecer informação pessoal online faz cada vez mais parte da vida moderna” e que “Não há alternativa para obter produtos ou serviços sem que tenha de fornecer dados pessoais.”.

Tabela 12- Opinião dos inquiridos acerca de 5 afirmações relacionadas com os seus sentimentos relativamente ao terem de fornecer informações pessoais online.

Nível de concordância	(1) Discordo Totalmente	(2) Discordo Parcialmente	(3) Indiferente	(4) Concordo Parcialmente	(5) Concordo Totalmente	Média	Mediana	Moda
Fornecer informação pessoal online faz cada vez mais parte da vida moderna	16 (10%)	23 (14,23%)	44 (27,50%)	44 (27,50%)	33 (20,63%)	3.34	3	4
Sinto que tenho de fornecer cada vez mais informação pessoal	11 (6,88%)	20 (12,50%)	37 (23,13%)	54 (33,75%)	38 (23,75%)	3.55	4	4
Sinto-me invadido ao fornecer informação pessoal.	2 (1,25%)	19 (11,88%)	50 (31,25%)	41 (25,63%)	48 (30%)	3.71	4	3
Importo-me de fornecer informação pessoal em troca de serviços online gratuitos, como o email	13 (8,13 %)	32 (20%)	42 (26,25%)	36 (22,50%)	37 (23,13%)	3.33	3	3
Não há alternativa para obter produtos ou serviços sem que tenha de fornecer dados pessoais.	15 (9,38%)	20 (12,50%)	51 (31,88%)	42 (26,25%)	32 (20%)	3.35	3	3

4.1.6. Comentários adicionais

Na última questão do questionário, houve 5 (cinco) inquiridos que apresentaram comentários adicionais.

Pela análise ao conteúdo das respostas verifica-se que estes inquiridos mostram ter noção de que os seus dados são constantemente recolhidos, e que as suas ações são constantemente monitorizadas:

“Dados pessoais são importantes para melhorar a experiência do utilizador. Por vezes podem ser utilizados por parte do *site*/plataforma para lucro o que a meu ver está correto desde que o utilizador o permita sendo este devidamente notificado. Sendo esta uma boa fonte de rendimento para os *sites* que permite ao utilizador não pagar para os utilizar acho que é uma boa solução desde que seja feita com total conhecimento de todos os envolvidos.” **(Inquirido nº15)**

“Infelizmente a nossa vida é cada vez mais monitorizada: quem somos, onde estamos, reconhecimento facial, impressão digital.” **(Inquirido nº73)**

No entanto, aceitam isso como uma consequência dos serviços a que querem aceder, tendo a ideia de que não têm poder de escolha sobre a recolha dos seus dados:

“O facto de ter consciência dos perigos da entrega dos dados não impede que o faça, porque, caso contrário, estou limitado praticamente a todos os níveis sociais” **(Inquirido nº112)**

“Mesmo que digamos expressamente que não queremos que os nossos dados sejam recolhidos e utilizados para outros fins, sou da opinião que na prática os usam na mesma.” **(Inquirido nº129)**

Há ainda um inquirido que considera que os utilizadores online cada vez se preocupam menos com a sua privacidade. No entanto esta despreocupação pode-se dever ao facto de estes não terem conhecimento do tratamento dado aos seus dados.

“Na minha opinião, cada vez mais, as pessoas preocupam-se menos com a sua privacidade.” **(Inquirido nº138)**

4.2. Análise dos *cookies* em *sites* de comércio eletrónico

Ao longo desta secção serão explanadas as Políticas de Privacidade de cada *site*, as tipologias e objetivos dos *cookies* presentes nesses mesmos *sites* e, por fim, é apresentada uma lista dos principais *cookies* usados por cada *site* e a sua função.

4.2.1. Políticas de privacidade dos *sites* de comércio eletrónico

Começamos por apresentar os aspetos relacionados com os *sites* retalhistas: amazon.com, continente.pt, decathlon.com, fnac.pt, sportzone.pt, wook.pt e Worten.pt.

amazon.com

Na primeira vez que se acede ao *site* não é apresentada nenhuma mensagem relativa à utilização de *cookies*. No entanto temos, no separador dedicado à ajuda ao cliente²⁰, várias questões sobre o tratamento de dados dos utilizadores online. A *amazon.com* explica apenas que utiliza a informação pessoal dos utilizadores para:

- Compra e entrega de produtos e serviços.
- Fornecer, solucionar problemas e melhorar os serviços da Amazon.
- Recomendações e personalização.
- Fornecer serviços de voz, imagem e câmara.
- Cumprir as obrigações legais.
- Comunicar com o cliente.
- Publicidade.
- Prevenção de Fraudes e Riscos de Crédito.

Quanto aos *cookies* utilizados, a Amazon²¹ explica que os utiliza para reconhecer o navegador do utilizador ou o tipo de dispositivo, aprender mais sobre os seus interesses e fornecer características e serviços para fins adicionais, tal como:

- Reconhecer o utilizador quando este se inscreve para utilizar os serviços da Amazon;
- Manter um registo dos itens guardados no carrinho de compras;

²⁰https://www.amazon.com/gp/help/customer/display.html?ie=UTF8&nodeId=468496&ref_=footer_privacy#GUID-1B2BDAD4-7ACF-4D7A-8608-CBA6EA897FD3_SECTION_7CE768C1281346DAAF87CE795E8B3610 [Consultado em 12/05/2020]

²¹ <https://www.amazon.com/gp/help/customer/display.html/?nodeId=201890250> [Consultado em 12/05/2020]

- Conduzir pesquisas e diagnósticos para melhorar o conteúdo, produtos e serviços da Amazon;
- Prevenir atividades fraudulentas e melhorar a segurança;
- Entrega de conteúdos, incluindo anúncios relevantes para os interesses do utilizador em *sites* da Amazon e *sites* de terceiros;
- Relatórios que permitem medir e analisar o desempenho dos serviços.

A Amazon não detalha que tipos de *cookies* utiliza nem fornece nenhuma opção relativa à desativação de alguns *cookies* ou bloqueio do rastreamento da atividade do utilizador *online*.

continente.pt

No primeiro acesso ao *site* do Continente, é apresentada uma mensagem onde o utilizador é informado da utilização de *cookies* por parte do *site*, mas não tem a opção de aceitar ou rejeitar.

O Continente não apresenta quais os tipos de dados do utilizador que recolhe, no entanto, ao aceder ao *site*²², podemos ver que na Política de *Cookies* o Continente informa que utiliza:

- ***Cookies de Sessão***, para analisar padrões de tráfego do *site*, permitindo-lhes melhorar o conteúdo e melhorar a usabilidade do *site*;
- ***Cookies de Análise***, para quantificar o número de utilizadores e realizar a medição e análise estatística de como é que os utilizadores usam o serviço oferecido, com o objetivo de melhorar o fornecimento de produtos ou serviços;
- ***Third Party Cookies***, para gerir a oferta de espaços publicitários que existem no *site*, analisando os hábitos de navegação do utilizador mostrando publicidade relacionada com o seu perfil de navegação.

É de salientar que o acesso à política de *cookies* não está visível no *site*, sendo apenas possível aceder com uma pesquisa a partir da Google com a entrada “políticas de *cookies* continente”.

Visto que é apresentada a mensagem referente à utilização dos *cookies*, os utilizadores podem *clicar* em “Saber mais”, onde é apresentado um conjunto de questões, e

²² <https://www.continente.pt/stores/continente/pt-pt/public/Pages/Politica-de-Cookies.aspx> [Consultado em 13/05/2020]

apresentam as instruções para o utilizador alterar a sua privacidade de navegação no navegador que está a utilizar.

decathlon.com

No primeiro acesso ao *site* da Decathlon, não é apresentada uma mensagem onde o utilizador é informado da utilização de *cookies* por parte do *site*.

A Decathlon não informa que tipos de dados recolhe nem que tipos de *cookies* utiliza, mas informa que recolhem²³ informações pessoais de diversas formas:

- Numa visita ao *site*, recolhem informação sobre o dispositivo. Esta informação é automaticamente fornecida pelo browser (tipo de dispositivo, versão do sistema operativo, identificador único do dispositivo, tipo e versão do browser e informação da rede móvel) ou informação de registo (endereço IP, identificador do telefone inteligente, resolução do ecrã e domínio de referência).
- Se se tornar um utilizador registado - Exigem que os utilizadores enviem informações de contato se quiserem comprar produtos ou serviços do *site* ou receber a newsletter.
- Se adquirir um produto ou serviço a partir do *site* - recolhem informações de contacto, tais como nome, número de telefone, endereço de e-mail, informações de pagamento, e endereço de envio.
- A partir de *cookies* e outras tecnologias, podem recolher informações tais como a atividade online do utilizador, as páginas do *site* que o utilizador visitou, os termos de pesquisa utilizados, e qualquer link ou publicidade em que o utilizador clique.

No *site* não é fornecida nenhuma opção relativa à desativação de alguns *cookies* ou bloqueio do rastreamento da atividade do utilizador *online*.

fnac.pt

No primeiro acesso ao *site* da Fnac, é apresentada uma mensagem onde o utilizador é informado da utilização de *cookies* por parte do *site*, tendo a opção de aceitar ou de configurar as opções do uso de *cookies*.

A Fnac não informa que tipos de dados recolhe nem de que forma, mas na Política de *Cookies*²⁴ informa que utiliza *cookies* criados pela própria Fnac para fins de navegação

²³ <https://www.decathlon.com/pages/privacy-policy> [Consultado em 14/05/2020]

²⁴ <https://www.fnac.pt/PoliticaCookies> [Consultado em 15/05/2020]

no *site*, otimização e personalização dos serviços. Utiliza também *Third Party Cookies* para identificar os interesses do utilizador através dos produtos consultados ou comprados e personalizar a oferta de publicidade, não tendo controlo sobre estes.

A Fnac explica que utiliza os seguintes *cookies*:

- Estritamente Necessários, que permitem ao utilizador aceder aos principais recursos do *site* (o uso do carrinho de compras ou o acesso à conta). Sem estes *cookies*, o *site* não pode funcionar normalmente e são criados pela Fnac.
- De Análise/ Performance, que permitem conhecer o uso e desempenho do *site* e melhorar o seu funcionamento (pesquisas feitas, frequência das páginas visualizadas)
- Funcionais, que permitem personalizar a experiência do utilizador (para relembrar ao utilizador os últimos produtos consultados, memorizar os itens no carrinho antes de continuar as suas compras).
- Publicidade, que servem para apresentar anúncios ou enviar informações adaptadas aos interesses do utilizador no *site* ou fora do nosso *site* durante a sua navegação na Internet. Ajudam a medir a eficácia de uma campanha publicitária e são usados para limitar o número de vezes que um anúncio é visualizado.

É também apresentada uma lista dos *cookies* utilizados, mencionados na secção 4.2.3.. Visto que os utilizadores têm a opção de configurar os *cookies*, a Fnac fornece a opção de os utilizadores controlarem os seus dados.

sportzone.pt

Ao aceder ao *site* da SportZone, aparece uma mensagem a informar que utilizam *cookies* próprios e de terceiros para melhorar a experiência dos utilizadores e os serviços do *site*. Ao *clicar* em qualquer parte do *site* a utilização de *cookies* é automaticamente aceite. Se se quiser configurar as opções do uso de *cookies*, é apresentada a opção na mensagem.

A SportZone não informa que tipos de dados recolhe nem de que forma, mas ao aceder à Política de *Cookies*²⁵, discriminam quais os *cookies* essenciais e não essenciais que utilizam, bem como uma lista dos mesmos (mencionados na secção 4.2.3.).

Desta forma, segundo a SportZone, estão inseridos no *site* *cookies*:

²⁵ <https://www.sprintersports.com/pt/enable-cookies> [Consultado em 16/05/2020]

- Estritamente necessários, para o funcionamento do *site*. Incluem *cookies* que permitem fazer login em áreas seguras dos *sites* ou fazer compras.
- De análise, que permitem reconhecer e contar o número de visitantes e ver como eles se movimentam no *site* durante o seu uso, ajudando a melhorar a maneira como o *site* funciona garantindo, por exemplo, que os utilizadores encontram facilmente o que procuram. Utilizam ainda *cookies* para compilar estatísticas de visitantes, como o número de pessoas que visitaram o *site*, como chegaram a o *site*, que tipo de tecnologia usam, quanto tempo passam no *site* e a que páginas acedem.
- De funcionalidade, para reconhecer o utilizador quando voltar ao *site*, permitindo personalizar o conteúdo e lembrar as preferências do utilizador

Quanto aos *cookies* de análise, a SportZone afirma que essas informações serão usadas pelos mesmos e por terceiros para tornar os *sites* e a publicidade neles exibida mais relevantes para os interesses do utilizador. Podem também partilhar essas informações com terceiros para essa finalidade, mas apenas se o utilizador consentir essa partilha.

Visto que os utilizadores têm a opção de configurar os *cookies*, a SportZone fornece a opção de os utilizadores controlarem os seus dados.

wook.pt

Da mesma forma que na SportZone, a Wook apresenta, ao aceder ao *site*, uma mensagem, a informar a utilização de *cookies*, onde a contínua navegação faz com que essa utilização seja automaticamente aceite. Se se quiserem configurar as opções do uso de *cookies*, é apresentada a opção na mensagem.

A Wook não apresenta que dados recolhe, mas na Política de *Cookies*²⁶ da **wook.pt**, é explicado o tratamento dado aos dados pessoais utilizadores. Desta forma, a **wook.pt** recolhe os dados dos utilizadores para as seguintes finalidades:

- Finalizar uma compra (recolhe o endereço de e-mail, morada de entrega e número de telefone) ou indicar o endereço de entrega (que fica guardado para futuras compras)
- Caso aceite o envio de mensagens comerciais por correio eletrónico ou SMS, aquando registo no *site*

²⁶ <https://www.wook.pt/Mensagem/index/?key=politicaprivacidade> [Consultado em 16/05/2020]

- Envio de newsletters, inquéritos de opinião, ou outras informações sobre produtos requisitados ou comprados pelo utilizador;
- Garantir que o *Site* vai ao encontro das necessidades do utilizador, através do desenvolvimento e publicação de conteúdos adaptados ao tipo de utilizador, da melhoria das capacidades de pesquisa e funcionalidades do *site* e da obtenção de informação estatística relativamente ao perfil tipo do utilizador;

A Wook esclarece que não partilha dados dos seus utilizadores com terceiros sem o consentimento do utilizador e apresenta os tipos de *cookies* utiliza (mencionados na secção 4.2.3.).

Visto que os utilizadores têm a opção de configurar os *cookies*, a Wook fornece a opção de os utilizadores controlarem os seus dados.

worten.pt

A Worten também apresenta, ao aceder ao *site*, uma mensagem, a informar a utilização de *cookies*, onde a contínua navegação faz com que essa utilização seja consentida. Se se quiserem configurar as opções do uso de *cookies*, é apresentada a opção na mensagem.

A Worten, na sua Política de Privacidade²⁷, afirma que trata dados de identificação, dados de contacto e morada do cliente, histórico de compras, perfil do cliente e comunicação de campanhas e novidades. No entanto não explica como recolhe estes dados.

A Worten explica ainda que utiliza os seguintes *cookies*:

- *First Party Cookies* de Sessão: servem para melhorar o conteúdo e melhorar a usabilidade do *site*. A informação recolhida por estes *cookies* serve para analisar padrões de tráfego do *site*. São *cookies* temporárias que permanecem em arquivo de *cookies* do navegador até que este abandone a página web, pelo que não fica qualquer registo no disco rígido do utilizador.
- *Cookies* de Análise/Performance: tratados pela Worten ou por terceiros, permitem analisar a navegação do utilizador no *site*, com o objetivo de melhorar o fornecimento de produtos ou serviços que lhe disponibilizamos.
- *Cookies* publicitárias: tratados pela Worten ou por terceiros, permitem gerir de uma forma mais eficaz a oferta de espaços publicitários que existem no *site*,

²⁷ <https://www.worten.pt/politica-de-privacidade> [Consultado em 18/05/2020]

podendo analisar os hábitos de navegação do utilizador e mostrar-lhe publicidade relacionada com o seu perfil de navegação.

Visto que os utilizadores têm a opção de configurar os *cookies*, a Worten fornece a opção de os utilizadores controlarem os seus dados

Passando para conjunto de *sites* de **C2C**, são apresentados os principais aspetos relativos aos *sites*: *custojusto.pt* e *ebay.com*.

custojusto.pt

Ao aceder ao *site*, o CustoJusto apresenta uma mensagem a informar a utilização de *cookies*, com a opção de aceitar ou rejeitar.

A Política de Privacidade²⁸ do CustoJusto informa que os dados pessoais que recolhem incluem informações como moradas pessoais, números de telemóvel, nomes de utilizador endereço de correio eletrónico, fotografias de perfis, preferências pessoais e hábitos de compra e conteúdos gerados pelos utilizadores. Pode também incluir identificadores numéricos únicos como o endereço IP do computador ou o endereço do aparelho móvel. No caso de criar uma conta profissional, são também recolhidos dados obrigatórios como nome da empresa, número fiscal, categoria profissional e número de Licença de Mediação Imobiliária.

A partir da navegação nas suas páginas web, o CustoJusto recolhe informação sobre o dispositivo (se é um telemóvel, tablet, computador), o seu modelo, sistema operativo, operador móvel, tipo de acesso à internet, fuso horário e o endereço de IP. Corre também dados de atividade, tendo assim informação das páginas que o utilizador visitou ou pesquisou, quando e se fez o login.

O CustoJusto também recolhe informações junto de outras fontes, tais como parceiros, clientes, anunciantes, fornecedores ou fontes públicas (como por exemplo do Turismo de Portugal ou do IMPIC).

Quanto aos *cookies* criados pelo CustoJusto, estes têm finalidades analíticas e técnicas. As analíticas servem para analisar o tráfego das páginas anonimamente e número de visitantes, sem armazenar dados pessoais, recolhendo assim informações estatísticas sobre o uso dos serviços pelos utilizadores. As técnicas, são *cookies* essenciais, utilizados

²⁸ <https://www.custojusto.pt/ajuda-rules?section=privacidade>

para fins relacionados com a prestação de serviços e para melhorar a navegação no *site* ou memorizar pesquisas.

Relativamente aos *cookies* instalados por terceiros, CustoJusto informa que servem para publicidade personalizada (através de funcionalidades da Google) e para finalidades analíticas, a partir de ferramentas Xiti²⁹.

Visto que os utilizadores têm a opção de configurar os *cookies*, o CustoJusto fornece a opção de os utilizadores controlarem os seus dados.

ebay.com

Visto que o RGPD se aplica a empresas ou entidades cujo tratamento de dados pessoais seja efetuado numa sucursal estabelecida na UE ou a empresas fora da UE que ofereçam serviços e bens ou controla os dados pessoais dos utilizadores na UE, o Ebay tem como obrigação fornecer, aos cidadãos da UE, as informações e proteção exigidas pelo RGPD.

É o que acontece quando abrimos ebay.com e nos aparece a mensagem para se aceitar os *cookies*. Ao clicar em “Mais informações” aparece o aviso de que o conteúdo e as preferências relativas aos *cookies* que os utilizadores querem ou não ativar só afetam clientes na Área Económica Europeia.

Na página “Publicidade e preferências relacionadas”³⁰, o utilizador pode controlar as informações que o eBay utiliza para mostrar publicidade. Informa também que partilha os dados dos utilizadores diretamente com o Facebook e Facebook Pixel para a publicidade nas plataformas do Facebook (nesta secção será também abrangido o tratamento de dados pessoais do Facebook).

O Ebay recolhe informações relativas ao tipo de navegador e as suas configurações, tipo de dispositivo e sistema operacional, informações sobre outros identificadores atribuídos ao dispositivo, segmentação do utilizador, endereço de IP, localização geográfica do dispositivo, endereço de email e informações relativas à atividade do utilizador no *site*.

O utilizador também tem a opção de desativar os seguintes tipos de recolha de dados, tendo também acesso a todas as empresas parceiras de cada tipo de recolha de dados:

²⁹ XiTi é uma ferramenta de análise web fornecida pela AT Internet para uma introdução à medição do tráfego web. Esta ferramenta está disponível em <https://www.xiti.com/en/>

³⁰ <https://www.ebay.com/gdpr> [Consultado em 19/05/2020]

- Personalização: significa que são recolhidos os dados pessoais do utilizador de modo a posteriormente personalizar a publicidade e/ou o conteúdo, adequando-os ao utilizador noutros contextos ao longo do tempo, como outros *sites*.
- Seleção, disponibilização e comunicação de anúncios: significa que os dados pessoais do utilizador recolhidos servem para selecionar anúncios, tratar dados sobre os anúncios publicitários que foram apresentados, a frequência com que foram apresentados, quando e onde foram apresentados e saber se o utilizador reagiu de alguma forma à publicidade (se clicou num anúncio ou efetuou uma compra).
- Conservação e acesso a informações: significa que recolhe informações já conservados no dispositivo do utilizador, como identificadores de publicidade, identificadores de dispositivos, *cookies* e tecnologias semelhantes.
- Publicidade na Google: significa que a Google coloca *cookies* e coleta informações sobre o seu dispositivo e uso deste *site*. (nesta secção será também abrangido o tratamento de dados pessoais da Google)
- Seleção, disponibilização e comunicação de conteúdo: significa que utilizam informações recolhidas sobre os interesses do utilizador, de modo a selecionar conteúdo, tratar dados sobre o conteúdo que foi apresentado, a frequência ou durante quanto tempo o mesmo foi apresentado, quando e onde foi apresentado e saber se reagiu de alguma forma ao conteúdo.

O Ebay não informa que tipos de *cookies* utiliza.

Foram também analisadas duas redes sociais, *facebook.com* e *linkedin.com*. Apresentamos abaixo os principais aspetos.

facebook.com

Ao aceder ao *site* do Facebook aparece uma mensagem a informar que utilizam *cookies* próprios e de terceiros para melhorar a experiência dos utilizadores e os serviços do *site*. Ao continuar a navegação a utilização de *cookies* é automaticamente aceite. Se se quiser configurar as opções do uso de *cookies*, é apresentada a hiperligação para a política de *cookies*, onde é possível fazer alterações.

Na página da Política de Dados³¹ do Facebook, é informado que este recolhe dados relativos a tudo o que os utilizadores fazem e fornecem, o que inclui informações sobre quando se cria uma conta, se cria ou partilha conteúdos ou se envia mensagens a outras pessoas. Estas ações incluem a recolha de dados sobre a data em que um ficheiro foi criado ou a localização de uma foto. No entanto há dados que estão abrangidos com proteções especiais. São estes as informações que o utilizador fornece voluntariamente sobre eventos de vida, campos do perfil do Facebook, saúde, estado civil ou o sexo que se tem interesse, ideologias políticas, saúde, origem racial ou étnica, filiação sindical e crenças filosóficas, visto que estão sujeitas a proteções especiais ao abrigo da legislação da UE.

O Facebook recolhe também informação sobre:

- tipos de conteúdos vistos pelo utilizador ou com os quais interage; as funcionalidades que utiliza; as ações que efetua; as pessoas ou as contas com as quais interage; e o tempo, a frequência e a duração das tuas atividades.
- as pessoas, páginas, hashtags e grupos a que os utilizadores estão ligados e sobre a forma como interagem. Se se importar ou sincronizar a lista de contactos, o Facebook recolhe esses dados para cruzamento com outros utilizadores, de forma a se encontrarem.
- Informações sobre transações efetuadas nos produtos do Facebook, como quando se compra algo num jogo ou se faz um donativo, o Facebook recolhe informações sobre a compra ou transação (inclui informações de pagamento, como o número do cartão de crédito/débito, detalhes de faturação, envio e contacto)

A partir do tipo de dispositivo, recolhem também informações sobre:

- Informações sobre os comportamentos e as operações efetuadas no dispositivo (se a janela está em primeiro ou segundo plano, ou os movimentos do rato (que podem ajudar a distinguir humanos de *bots*));
- Dados sobre o sistema operativo, versões do software e hardware, tipo de browser, força do sinal, nível de bateria, espaço de armazenamento disponível, nomes e tipos de apps, acesso ao GPS, à localização ou camara e fotografias;
- Nome da operadora móvel, idioma, fuso horário, número de telemóvel, endereço IP, velocidade da ligação

³¹ <https://www.facebook.com/about/privacy> [Consultado em 20/05/2020]

- Sinais de Bluetooth e informações sobre torres de rede móvel, *beacons* e pontos de acesso Wi-Fi nas proximidades;
- Dados de *cookies* armazenados no dispositivo, incluindo definições e identificações de *cookies*

Relativamente aos Dados de Armazenamento de *Cookies*³², estes são utilizados para os seguintes fins:

- Autenticação: para manter a sessão iniciada enquanto o utilizador navega entre páginas do Facebook, para lembrar o browser, para não se ter de introduzir os dados de início de sessão no Facebook continuamente e se poder iniciar sessão mais facilmente no Facebook através de apps e *sites* de terceiros.
- Segurança e integridade do *site* e de produtos: para ajudar a manter a segurança da conta, dados e dos Produtos do Facebook, para armazenar informações que permitem recuperar a conta quando o utilizador se esquece da palavra-passe ou para solicitar um processo de autenticação adicional quando a conta foi pirateada. Também são utilizados *cookies* para combater atividades que desrespeitam as políticas do Facebook ou que comprometem a capacidade de fornecer os Produtos do Facebook (lutar contra o spam e ataques de *phishing* ao permitir identificar computadores que são utilizados para criar grandes quantidades de contas do Facebook falsas)
- Funcionalidades e serviços do *site*: armazenar preferências, saber quando se viu ou interagiu com um conteúdo e fornecer conteúdos personalizados. Armazenam também informações num cookie que é instalado no dispositivo para fornecer conteúdos relevantes face à localização (como ver um *site* no idioma certo)
- Desempenho: perceber a velocidade de carregamento dos conteúdos e encaminhar o tráfego entre os servidores, registar a proporção do ecrã
- Estatísticas e pesquisa: analisar como é que o utilizador utiliza os serviços, que partes do Facebook são mais relevantes e identificar funcionalidades que possam ser melhoradas
- Publicidade, recomendações, estatísticas e medição: apresentar anúncios e recomendar negócios ou organizações do interesse do utilizador, medir o

³² <https://www.facebook.com/policies/cookies/> [Consultado em 21/05/2020]

desempenho de campanhas de anúncios dos negócios ou avaliar a quantidade de vezes que o anúncio aparece ao utilizador (Facebook, 2020).

É possível desativar o Rastreo de Atividade Fora do Facebook em https://www.facebook.com/off_facebook_activity/

linkedin.com

Tal como o Facebook, no *site* do LinkedIn aparece uma mensagem a informar que utilizam *cookies* para melhorar o serviço e exibir anúncios personalizados . Ao continuar a navegação a utilização de *cookies* é automaticamente aceite. Não é apresentada a opção de configurar as opções do uso de *cookies*, é apenas apresentada a hiperligação para a política de *cookies*, onde são esclarecidas algumas questões.

O LinkedIn não menciona que tipos de dados recolhe nem como, mas na sua Política de *Cookies*³³, que é bastante resumida, descreve apenas que a finalidade desta tecnologia serve para:

- Autenticação: para reconhecer o utilizador quando este acede à conta
- Segurança: para ativar e dar suporte aos recursos de segurança, manter a conta segura e para ajudar a detetar atividades nocivas e violações do nosso Contrato do Utilizador.
- Preferências, recursos e serviços: para ajudar a preencher formulários, ter informações e conteúdos personalizados e ara lembrar informações sobre o navegador e preferências
- Conteúdo personalizado: para lembrar pesquisas anteriores, podendo oferecer informações adicionais
- Plugins dentro e fora do LinkedIn: por exemplo, o botão “Candidate-se pelo LinkedIn” pode ser encontrado na rede social ou em *sites* de terceiros, reconhecendo o utilizador e encaminhando-o entre páginas. (LinkedIn, 2020)

Não é mencionada a opção de alterar a utilização de *cookies* por parte do LinkedIn.

Por fim, foi analisado o maior portal na *Internet*, a **google.com**.

Ao aceder à página principal da Google, não é apresentada nenhuma mensagem relativa ao uso de *cookies*.

³³ <https://www.linkedin.com/legal/cookie-policy> [Consultado em 21/05/2020]

A sua Política de Privacidade ³⁴ indica que a Google recolhe informações como o idioma, aplicações, navegadores e suas definições, dispositivos (em Android para fazer atualizações automáticas ou reduzir a luminosidade do ecrã quando se tem pouca bateria), sistema operativo, informações da rede móvel (incluindo nome e número de telefone do operador), interação das aplicações, atividade do sistema e relatórios de falhas, localização (GPS, endereço IP, informações sobre itens próximos do dispositivo como pontos de acesso Wi-Fi ou torres de redes móveis).

Durante a atividade, recolhe informações sobre termos pesquisados, vídeos vistos, visualizações e interações com conteúdos, atividade de compra, informações de voz e atividade em *sites* ou aplicações de terceiros que utilizam os serviços da Google. Ao utilizar os serviços para receber ou fazer chamadas e enviar ou receber mensagens, recolhem informações sobre o registo telefónico (número do autor da chamada, do destinatário, hora, data e duração das chamadas).

Quanto às categorias de *cookies* utilizadas para executar os *sites* da Google e produtos relacionados com publicidade, explicam que são usadas para³⁵:

- Preferências: permitem que *sites* lembrem informações que mudam a forma como o *site* se comporta ou é exibido, como o idioma preferido ou a região em que o utilizador está
- Segurança: para autenticar os utilizadores, evitar a utilização fraudulenta de credenciais de login e proteger os dados do utilizador de terceiros não autorizados.
- Processos: ajudam a que o *site* funcione e forneça serviços que o visitante espera, como navegar em páginas da web ou aceder a áreas seguras do *site*. Sem esses *cookies*, o *site* pode não funcionar corretamente.
- Publicidade: para selecionar publicidade com base no que é relevante para o utilizador, melhorar os relatórios sobre o desempenho da campanha e evitar a exibição de anúncios que o utilizador já viu, lembrar as pesquisas mais recentes do utilizador, as interações anteriores com anúncios. São utilizados *cookies* de conversão com o objetivo principal de ajudar anunciantes a definir quantas vezes as pessoas que clicam nos anúncios acabam por se direcionar para o *site* publicitado. Estes *cookies* são maioritariamente utilizados por terceiros.

³⁴ <https://policies.google.com/privacy?hl=pt-PT#infocollect> [Consultado em 21/05/2020]

³⁵ <https://policies.google.com/technologies/types?hl=pt-BR> [Consultado em 22/05/2020]

- Google Analytics: é a ferramenta de análise da Google que ajuda os proprietários de *sites* e apps a entender como é que os utilizadores se envolvem com os serviços. Usa um conjunto de *cookies* para recolher informações e informar estatísticas de uso do *site* sem identificar cada visitante para a Google.

Não é mencionada a opção de alterar a utilização de *cookies* por parte da Google.

4.2.2. Tipo de *cookies* presentes nos *sites* de comércio eletrónico

Nesta secção são apresentados, através dos dados secundários recolhidas pela ferramenta de pesquisa da Cookiepedia, os resultados obtidos relativamente à tipologia de *cookies* utilizados em cada *site* e os respetivos objetivos.

Na Tabela 13 vê-se os resultados obtidos quanto à tipologia de *cookies* utilizados nos 12 *sites*. Os *Third Party Cookies* são o ecossistema de tecnologia de anúncios, sendo possível verificar que são os que estão mais presentes nos *sites*. Os *Cookies* Persistentes também são dos mais utilizados, visto que para as empresas é uma mais valia que estes se mantenham nos dispositivos dos utilizadores *online*. Os *Cookies* de Sessão são os menos presentes, talvez porque estes tipos de *cookies* apenas permanecem num navegador e retêm as informações do utilizador até que ele seja fechado. Na tabela verifica-se também que a Google e o Facebook são os *sites* onde foi detetada a maior existência de *cookies*, ao contrário da Wook e Decathlon, que apresentam uma quantidade baixa.

Tabela 13- Tipos de *Cookies* utilizados pelos *sites* de comércio eletrônico selecionados

<i>Site</i>	Total de cookies	<i>First Party-Cookies</i> Persistentes	<i>First Party-Cookies</i> de Sessão	<i>Third Party Cookies</i>
amazon.com	1946	1,90%	3,70%	94,40%
continente.pt	94	37,23%	11,70%	51,06%
decathlon.com	52	46,15%	5,77%	48,08%
fnac.pt	113	46,90%	13,27%	39,82%
sportzone.pt	146	25,34%	2,05%	72,60%
wook.pt	43	30,23%	11,63%	58,14%
worten.pt	146	40,41%	3,42%	56,16%
custojusto.pt	250	6,80%	32,40%	60,80%
ebay.com	4254	18,19%	6,23%	75,58%
facebook.com	18018	1,40%	1,55%	97,04%
linkedin.com	3793	6,75%	9,46%	83,79%
google.com	24764	6,70%	4,16%	89,14%

Categorizando estes dados por objetivos, obtiveram-se os resultados apresentados na Tabela 14. Visto que, como se verifica acima, os *Third Party Cookies* têm uma grande presença nos *sites*, é de esperar as grandes percentagens de *cookies* com objetivo publicitário, sendo a publicidade *online* a utilização mais comum de *Third Party Cookies*. Há uma percentagem muito baixa de *cookies Estritamente Necessários* e de *Funcionalidade*, com percentagens inferiores a 10%, com uma única exceção (os 15% do eBay nos de funcionalidade. Observamos também que é bastante elevada a percentagem de *cookies Desconhecidos*, principalmente na Google, com 83,22% de *cookies Desconhecidos*. Com exceção da Wook, os *cookies* de *Publicidade* e os *Desconhecidos* são os que mais existem estão em primeiro e segundo lugar na maioria dos *sites*. Apenas em três *sites* (Fnac, LinkedIn e Google) os *cookies* de *Desconhecidos* estão em primeiro lugar e os de *Publicidade* em segundo.

Tabela 14- Objetivo dos *Cookies* utilizados pelos *sites* de comércio eletrônico selecionados

<i>Site</i>	Análise/ Performance	Estritamente Necessários	Funcionalidade	Publicidade	Desconhecidos
amazon.com	0,67%	0,05%	1,03%	80,58%	17,68%
continente.pt	12,77%	4,26%	4,26%	52,13%	26,60%
decathlon.com	21,15%	7,69%	0%	28,85%	42,31%
fnac.pt	19,47%	1,77%	1,77%	32,74%	44,25%
sportzone.pt	8,90%	2,74%	0,68%	60,27%	27,40%
wook.pt	23,26%	4,65%	6,98%	48,84%	16,28%
worten.pt	5,48%	3,42%	2,74%	55,48%	32,88%
custojusto.pt	1,20%	0,40%	0,40%	77,20%	20,80%
ebay.com	6,23%	1,50%	15,63%	55,48%	32,88%
facebook.com	4,79%	5,29%	2,50%	52,64%	34,78%
linkedin.com	3,43%	3,03%	7,28%	31,82%	54,44%
google.com	5,12%	0,53%	1,08%	10,05%	83,22%

4.2.3. Cookies presentes nos sites de comércio eletrônico

Nas tabelas 15 a 21 elaboradas com o propósito de se perceber qual a verdadeira utilidade de um *cookie* num *site*. Assim, apresentamos os nomes dos *cookies*, o seu tipo e objetivo e qual a função em cada *site*.

continente.pt

O continente.pt apresenta uma lista de *cookies* que utiliza no seu *site*. Apresentamos abaixo, na Tabela 15 alguns exemplos dos principais *cookies* (ver lista completa no Anexo I) indicados pelo *continente.pt*, acrescentando o seu tipo e objetivo, de acordo com a Cookiepedia.

Tabela 15- Principais *cookies* utilizados pelo continente.pt

Cookie	Tipo de Cookie/ Objetivo	Função
AnonymousBasket	<i>First Party Cookies</i> Persistentes; Desconhecido	Mantém o registo dos produtos adicionados ao carrinho por um utilizador anónimo.
cPrompt_useCookies	<i>First Party Cookies</i> Persistentes; Desconhecido	Permite saber se o utilizador aceitou <i>cookies</i> no <i>site</i> .
JSESSIONID	<i>Third Party Cookies</i> ; Análise/ Performance	Preserva os estados dos utilizadores entre páginas.
AddThis	<i>First Party Cookies</i> Persistentes; Desconhecido	Serviço que permite a partilha de páginas por parte do utilizador em redes sociais.
E-goi	<i>Third Party Cookies</i> ; Desconhecido	Permite comunicar de forma personalizada, através de SMS e Email Marketing, de acordo com a utilização do <i>site</i> por parte do utilizador.
Google Analytics	A Cookiepedia não reconhece o nome deste cookie.	Aplicação da Google que permite armazenar dados da navegação, por parte dos utilizadores, ao Continente Online.

Adaptado de Continente (2020)

decathlon.com

O *site* da Decathlon não apresenta nenhuma lista dos *cookies* utilizados. Dessa forma, consultamos a base de dados elaborada para este estudo a partir da Cookiepedia e apresentamos, na Tabela 16, alguns exemplos de *cookies* utilizados pela *decathlon.com*, descrevendo o seu tipo, objetivo e função, de acordo com a Cookiepedia.

Tabela 16- Exemplos de *cookies* utilizados pela decathlon.com

Cookie	Tipo de Cookie/ Objetivo	Função
_shopify_fs	<i>First Party Cookies</i> Persistentes; Análise/ Performance	Recolher informações sobre como os visitantes utilizam o <i>site</i> ³⁶ . Não recolhem informação que identifique um visitante.
Rlas3	<i>Third Party Cookies</i> ; Publicidade	Fornecer anúncios mais relevantes para o utilizador e para os seus interesses. Limitar o número de vezes que se vê um anúncio e ajudar a medir a eficácia da campanha publicitária.
secure_customer_sig	<i>First Party Cookies</i> Persistentes; Estritamente Necessário	Essenciais para que se possa movimentar no <i>site</i> e utilizar as suas funcionalidades. Sem estes, serviços que se solicitou, como carrinhos de compras ou faturação eletrónica, não podem ser fornecidos.

fnac.pt

A *fnac.pt*³⁷ apresenta alguns dos *cookies* que utiliza e a sua função no *site*. A Tabela 17 apresenta uma seleção desses *cookies*, onde foi acrescentado o seu tipo e objetivo, de acordo com a Cookiepedia.

³⁶ Páginas visitadas com mais frequência e se recebem mensagens de erro das páginas *web*.

³⁷ <https://www.fnac.pt/PoliticaCookies> [Consultado em 15/05/2020]

Tabela 17- Exemplos de *cookies* utilizados pela *fnac.pt*

Cookie	Tipo de Cookie/ Objetivo	Função
SID	<i>First Party</i> Cookie Persistente; Estritamente Necessário	Adicionar ao carrinho e efetuar a compra
UID	<i>First Party</i> Cookie Persistente; Análise/ Performance	Identificar o cliente
ASP.NET_SessionID	<i>First Party</i> Cookie Persistente; Estritamente Necessário	Identificador de sessão
LSI	<i>First Party</i> Cookie Persistente; Desconhecido	Lista dos últimos artigos visitados
NBART	<i>First Party</i> Cookie Persistente; Desconhecido	Número de artigos no carrinho

Adaptado de Fnac (2020)

sportzone.pt

A SportZone apresenta também uma lista de 29 *cookies*, sendo alegadamente todos os *cookies* que utilizam. A Cookiepedia detetou, neste *site*, o uso de 146 *cookies*, no entanto a última vez que a Política de *Cookies* da SportZone foi atualizada foi em 2018.

Como a *sportzone.pt*³⁸ apresenta uma lista de 29 *cookies*. A Cookiepedia detetou, neste *site*, o uso de 146 *cookies*, no entanto a última vez que a Política de *Cookies* da SportZone foi atualizada foi em 2018. Desta forma consultamos a base de dados elaborada para este estudo a partir da Cookiepedia e apresentamos, na Tabela 18, alguns exemplos de *cookies* utilizados pela *sportzone.pt*, descrevendo o seu tipo, objetivo e função, de acordo com a Cookiepedia, de forma a obtermos dados mais atualizados.

³⁸ <https://www.sprintersports.com/pt/enable-cookies> [Consultado em 16/05/2020]

Tabela 18- Exemplos de *cookies* utilizados pela *sportzone.pt*

Cookie	Tipo de cookie/ Objetivo	Função
_pk_id*	<i>First Party</i> Cookie Persistente; Análise/ Performance	Usado para ajudar os proprietários de <i>sites</i> a rastrear o comportamento dos visitantes e medir o desempenho do <i>site</i> .
_dc_gtm_UA-2367023-1	<i>First Party</i> Cookie Persistente; Estritamente Necessário	Carregar outros <i>scripts</i> e códigos numa página. Quando é utilizado, pode ser considerado como estritamente necessário visto que sem ele, outros <i>scripts</i> podem não funcionar corretamente.
cto_tld_test	<i>First Party</i> Cookie Persistente; Publicidade	Utilizado para identificar o visitante através de visitas e dispositivos. Permite que o <i>site</i> apresente ao visitante publicidade relevante. O serviço é fornecido por concentrações de publicidade de terceiros.
awxxxx	<i>First Party</i> Cookie Sessão; Publicidade	Utilizado para perceber de que <i>site</i> vem o utilizador quando clica num dos link .

wook.pt

A *wook.pt* apresenta alguns dos *cookies* que utiliza e a sua função no *site*. No entanto, como a última atualização foi em 2018, optou-se por criar a Tabela 19, que apresenta alguns exemplos de *cookies* utilizados pela *wook.pt*, descrevendo o seu tipo, objetivo e função, de acordo com a Cookiepedia.

Tabela 19- Cookies utilizados pela wook.pt

<i>Cookies</i>	Tipo de cookie/ Objetivo	Função
__utma ³⁹	<i>First Party Cookies</i> Persistentes; Funcionalidade	Permite aos proprietários de <i>sites</i> acompanhar o comportamento dos visitantes e medir o desempenho do <i>site</i> . Calcula as estatísticas de visitantes novos e retornados. O cookie é atualizado sempre que são enviados dados para o Google Analytics.
JSESSIONID	<i>First Party Cookies</i> Sessão; Análise/ Performance	Utilizado para manter uma sessão de utilizador anónima pelo servidor.
BIGipServerPROD_rp_wood_pool	<i>First Party Cookies</i> Sessão; Estritamente Necessário	Associado à gestão de sessões, para assegurar que os pedidos dos utilizadores são encaminhados de forma consistente para o servidor correto.
WookCartId	<i>First Party Cookies</i> Persistentes; Desconhecido	Controlo do cesto de compras do utilizador

worten.pt

A *worten.pt* não apresenta nenhuma lista dos *cookies* utilizados. Dessa forma, consultamos a base de dados elaborada para este estudo a partir da Cookiepedia e apresentamos, na Tabela 20, alguns exemplos de *cookies* utilizados pela *worten.pt*, descrevendo o seu tipo, objetivo e função, de acordo com a Cookiepedia.

³⁹ Pertence ao Google Analytics.

Tabela 20- *Cookies* utilizados em Worten.pt

Cookie	Tipo de cookie/ Objetivo	Função
_dc_gtm_UA-10378744-1	<i>First Party Cookies</i> Persistentes; Estritamente Necessário	Este cookie está associado a <i>sites</i> que utilizam o Google Tag Manager para carregar outros scripts e códigos numa página. Quando é utilizado, pode ser considerado como estritamente necessário visto que sem ele, outros scripts podem não funcionar corretamente.
_fbp	<i>First Party Cookies</i> Persistentes; Publicidade	Utilizado pelo Facebook para fornecer uma série de produtos publicitários, tais como licitações em tempo real de anunciantes terceiros
vuid	<i>First Party Cookies</i> Persistentes; Funcionalidade	Estes <i>cookies</i> são utilizados pelo leitor de vídeo Vimeo em <i>sites</i> .
_hjIncludedInSample	<i>First Party Cookies</i> Sessão; Performance	Associado à funcionalidade e serviços de análise web. Identifica de forma única um visitante durante uma única sessão do navegador e indica que ele está incluído numa amostra de audiência.

custojusto.pt

O custojusto.pt não apresenta nenhuma lista dos *cookies* utilizados. Dessa forma, consultamos a base de dados elaborada para este estudo a partir da Cookiepedia e apresentamos, na Tabela 21, alguns exemplos de *cookies* utilizados pelo **custojusto.pt**, descrevendo o seu tipo, objetivo e função, de acordo com a Cookiepedia.

Tabela 21- *Cookies* utilizados em *custojusto.pt*

Cookie	Tipo de cookie/ Objetivo	Função
__cfduid	First Party- <i>Cookies</i> Persistentes; Estritamente Necessário	Cookie associado a <i>sites</i> que utilizam o CloudFlare, utilizado para acelerar o tempo de carregamento de páginas. Utilizado para anular quaisquer restrições de segurança com base no endereço IP do utilizador.
atuserid	First Party- <i>Cookies</i> Persistentes; Análise/ Performance	Faz parte do sistema analítico do <i>site</i> para recolher informações sobre a atividade dos visitantes. Esta informação é enviada para empresas de análise e utilizada para melhorar os serviços que presta aos seus utilizadores.
GoogleAdServingTest	First Party- <i>Cookies</i> Sessão; Publicidade	Utilizado para determinar que anúncios foram mostrados ao visitante do <i>site</i> .
ad-privacy	First Party- <i>Cookies</i> Persistentes Publicidade	Geralmente fornecido pelo amazon-adsystem.com para fins publicitários.
cto_lwid	First Party- <i>Cookies</i> Persistentes; Funcionalidade	Mantém um motor de reservas em funcionamento.

CAPÍTULO V – DISCUSSÃO DOS RESULTADOS

Neste ponto da dissertação serão analisados e discutidos os resultados obtidos através do questionário *online* e da análise aos *cookies* num conjunto de *sites* de comércio eletrónico com base na revisão da literatura efetuada no Capítulo II.

Os resultados anteriormente apresentados visam avaliar o nível de consciencialização do utilizador em relação à partilha e recolha dos seus dados, percebendo o seu nível de conhecimento em relação ao conceito de dados pessoais, avaliando o seu comportamento perante a partilha dos seus dados, percebendo o seu nível de conhecimento perante os seus direitos inerentes no RGPD e perante as opções que os *sites* de comércio eletrónico oferecem para a sua utilização. Os resultados anteriores visam também definir quais são os dados recolhidos pelos *sites* de comércio eletrónico a partir de uma análise às políticas de privacidade dos *sites* e os tipos de *cookies* presentes nos mesmos.

Como se viu na secção 2.5.1.1 Informação recolhida pelos *sites*, de acordo com Laudon & Traver (2019), todas as opções apresentadas no questionário para a identificação de dados pessoais, são, de facto, IPI. Dada a tão baixa percentagem de inquiridos que selecionou a maioria das opções, os resultados obtidos mostram que os inquiridos não têm esta noção, mostrando que não estão informados quanto à natureza dos dados que fornecem. Sendo a maioria dos inquiridos jovens e com formação superior, este aspeto pode considerar-se surpreendente. Assim sendo, é necessária mais informação e sensibilização dos utilizadores online sobre estas questões e a sua importância. De qualquer das formas, a maioria dos inquiridos afirma fornecer os seus dados pessoais online em troca de algum benefício. As três principais razões para os inquiridos fornecerem estes dados são “Fazer um pagamento”, para “Conseguir a entrega de uma compra” e “Obter um serviço”.

O facto de os inquiridos terem o hábito de aceitar os *cookies* quando acedem a um *site*, e não tendo estes a noção de quais são os dados efetivamente pessoais, ou quais são os dados recolhidos pelos *sites* de comércio eletrónico, justifica a sensação de controlo *Parcial* sobre os seus dados, uma vez que têm a opção de, em alguns *sites*, aceitar os *cookies* ou configurá-los. Esta falsa sensação de quase controlo pode-se justificar visto que a maioria dos *sites* analisados apresentavam uma mensagem informativa sobre a utilização dos *cookies*, apresentando a opção de aceitar ou rejeitar.

Apesar da sensação de controlo *Parcial* sobre os seus dados, a percentagem de inquiridos que utiliza *software* que evita que as suas atividades sejam monitorizadas ou que tenham

mudado as definições de privacidade do navegador que utilizam é baixa, o que mostra que os utilizadores *online* não têm cuidado ou atenção sobre os dados que querem partilhar *online*. Uma grande percentagem de inquiridos evita alguns *sites* de comércio eletrónico por receio de que a sua atividade seja monitorizada, não tendo sido possível apurar o motivo, embora possamos admitir que se sentem desprotegidos ou sem capacidade para se protegerem de outra forma.

A maioria dos inquiridos afirma ter conhecimento sobre o RGPD, mostrando ter conhecimento sobre o Direito a Ser Esquecido (desenvolvido na secção 2.5.2.2.). No entanto, o mesmo não se aplica ao Direito à Portabilidade dos Dados (secção 2.5.2.1.) nem à aplicação do RGPD em *sites* de comércio eletrónico fora da UE. De acordo com o RGPD, uma pessoa pode solicitar ao responsável pelo tratamento de dados que apague os seus dados pessoais e com a portabilidade de dados e o utilizador *online* passa a poder exigir a uma empresa os dados que lhe dizem respeito num formato que permitirá a migração para outra empresa. O RGPD aplica-se a empresas ou entidades cujo tratamento de dados pessoais seja efetuado numa sucursal estabelecida na UE ou a empresas fora da UE que ofereçam serviços e bens ou controla os dados pessoais dos utilizadores na UE. Os inquiridos mostraram não ter conhecimento sobre estes três factos. Por exemplo, o Ebay, tem como obrigação, aos cidadãos da UE, fornecer as informações e proteção exigidas pelo RGPD. É o que acontece quando abrimos *ebay.com* e nos aparece a mensagem para se aceitar os *cookies*, ao contrário da *amazon.com*. O Ebay é bastante esclarecedor quanto ao tratamento dado aos dados pessoais dos utilizadores, apresentando as opções que o utilizador pode controlar para mostrar a publicidade no *site* e as empresas a quem partilha diretamente os dados dos utilizadores.

Percebemos que os inquiridos não concordam com nenhuma das condições que os *sites* de comércio eletrónico oferecem para a sua livre utilização. A maioria dos inquiridos não concorda em ter de pagar para não ser monitorizado enquanto navega num *site*, que os *sites* de comércio eletrónico partilhem os seus dados como utilizador sem a sua permissão, que utilizem os seus dados como utilizador para benefício próprio ou que a sua atividade *online* seja monitorizada em troca de acesso livre a um *site*. Ainda assim, acabam por aceitar os *cookies* sem uma revisão prévia, permitindo tudo isto de forma inadvertida, pois verifica-se que as organizações de comércio eletrónico não têm especial cuidado com este aspeto.

Os inquiridos sentem que têm de fornecer cada vez mais informação pessoal e que se sentem invadidos e mostram também que se importam de fornecer informação pessoal em troca de serviços online gratuitos, como o email. No entanto, acabam por reconhecer que fornecer informação pessoal faz cada vez mais parte da vida moderna e que não há alternativa para obter produtos ou serviços sem que tenha de fornecer dados pessoais. Estes dados permitiram perceber que os inquiridos apenas aceitam as condições que lhes são oferecidas pelos *sites* de comércio eletrónico relativamente à recolha dos seus dados porque não têm outra alternativa para aceder aos serviços, acabando por se tornar um hábito no seu dia a dia.

Percebemos que a maioria dos *sites* de comércio eletrónico, na sua política de privacidade, menciona que dados pessoais recolhem, e como, ou os tipos de *cookies* que utilizam. São poucos os que apresentam estes três pontos. Podemos afirmar que a maioria das empresas não cumprem o que está definido no RGPD.

Apesar de haver uma grande quantidade de *cookies* desconhecidos, os dados recolhidos dos utilizadores online são maioritariamente utilizados para motivos publicitários ou partilha com empresas terceiras, o que se justifica dada a grande quantidade de *Third Party Cookies*. Os *Cookies* Persistentes também são dos mais utilizados, visto que para as empresas é uma mais valia que estes se mantenham nos dispositivos dos utilizadores *online*. Os *Cookies* de Sessão são os menos presentes, talvez porque estes tipos de *cookies* apenas permanecem num navegador e retêm as informações do utilizador até que ele seja fechado. São poucos os *sites* que apresentam na sua Política de Privacidade os *cookies* que utilizam. A utilização de *cookies* Estritamente Necessários e de Funcionalidade é de tal forma baixa que mesmo não aceitando a utilização dos *cookies* num *site*, não se sentirá nenhuma dificuldade ao navegar no *site* (visto que estes *cookies* são importantes para a navegação e usabilidade do *site*).

No problema de pesquisa inicialmente proposto pretendia-se perceber a utilização dos Big Data no contexto do comércio eletrónico, quer do ponto de vista dos utilizadores, quer do ponto de vista das organizações. Devido à grande falha existente no estudo dos Big Data aplicado ao comércio eletrónico, sendo que este é bastante importante para uma organização, representando bastante valor, e à sua relação com a recolha de dados pessoais neste ambiente, este estudo contribui para que os utilizadores *online* tenham a perceção real de qual a utilização dos seus dados pessoais recolhidos em ambiente de comércio eletrónico e para as organizações adquirir conhecimento relativamente às preocupações dos utilizadores *online* quanto aos seus dados pessoais.

Para solucionar estes problemas, foi desenvolvido um Estudo de Caso com tipologia “caso múltiplo”. Um Estudo de Caso, segundo Yin (2001) é um método de investigação que é utilizado com muita frequência, principalmente quando a investigação tem como objetivo conhecer o “como?” e o “porquê?”. Neste caso, é múltiplo porque foi elaborada, primeiramente, uma revisão bibliográfica, para se adquirir conhecimento e base teórica para fundamentar os dois instrumentos de recolha de dados utilizados: um questionário *online* e uma análise a um conjunto selecionado de *sites* de comércio eletrónico.

Como contribuições académicas, nos procedimentos metodológicos desta pesquisa, foi realizada uma revisão teórica sobre os principais temas envolvidos, sendo desta forma possível entender o conceito de Big Data, identificar as suas características, principais tecnologias e benefícios ou desafios; definir quais as questões éticas e legais referentes aos Big Data e métodos de rastreamento *online* existentes.

A partir do questionário *online*, cujo objetivo era avaliar o nível de consciencialização do utilizador em relação à partilha e recolha dos seus dados em ambiente de comércio eletrónico, foram obtidas 160 respostas, onde foi possível apresentar características gerais da amostra por meio das respostas recebidas, apontar os dados que os utilizadores consideram ser pessoais e qual o motivo para a sua partilha para com os *sites* de comércio eletrónico, perceber quais os comportamentos dos utilizadores perante a sua noção de controlo sobre os seus dados, entender o nível do conhecimento do utilizador online em relação ao RGPD e compreender qual a posição dos utilizadores perante a recolha e utilização dos seus dados pessoais. Observamos que os inquiridos não têm noção de quais são de facto os dados pessoais que partilham, que a maioria dos inquiridos que tem o hábito de partilhar os seus dados pessoais têm o Ensino Superior, que fornecem os seus dados maioritariamente para fazer pagamentos, conseguir a entrega de uma compra ou

obter um serviço. Foi também possível perceber que apesar de os inquiridos sentirem pouco ou nenhum controlo sobre a informação que disponibilizam, têm o hábito de aceitar os *cookies* quando acedem a um *site*, que não utilizam *software* que evita o monitoramento da sua atividade nem que mudam as definições de privacidade do navegador que utilizam. No entanto, grande parte dos inquiridos evita aceder a alguns *sites* de comércio eletrónico com receio que a sua atividade seja monitorizada. Apesar de a maioria dos inquiridos considerar ter conhecimento sobre o RGPD, são poucos os que sabem de facto quais os seus direitos, não tendo noção de onde se aplica o RGPD e que têm o Direito à Portabilidade dos Dados. A maioria dos inquiridos sabe que têm o Direito a Ser Esquecidos. Por fim, este estudo revelou que, apesar dos inquiridos terem o hábito de partilhar os seus dados pessoais e de apenas aceitarem as condições que lhes são oferecidas pelos *sites* de comércio eletrónico relativamente à recolha dos seus dados, estes não concordam com nenhuma destas condições, sentindo que cada vez mais têm de fornecer informação pessoal e que se sentem invadidos.

Com a análise a um conjunto selecionado de *sites* de comércio eletrónico, foi possível analisar as suas Políticas de Privacidade e *cookies* utilizados em cada *site*, sendo possível a comparação entre aquilo que as empresas dizem ser o tratamento dados aos dados pessoais dos utilizadores, com o tratamento realmente dado. Este estudo apresenta, desta forma, as Políticas de Privacidade de todos os *sites* analisados e listas e resultados dos *cookies* utilizados, categorizados por tipo e objetivo. Neste estudo concluiu-se que a publicidade *online* é a utilização mais comum de *Third Party Cookies*. Ao adicionar as suas etiquetas a uma página, que pode ou não exibir anúncios, os anunciantes podem rastrear um utilizador ou um dispositivo em muitos dos *sites* que visitam. Percebemos também que a grande maioria dos *sites* analisados partilha os dados recolhidos com empresas terceiras. São poucos os *sites* que utilizam *cookies* como análise de dados ou funcionalidade do *site* e que os *cookies* estritamente necessários têm pouca presença nos *cookies* utilizados pelos *sites* de comércio eletrónico.

Quanto às limitações do nosso estudo, reconhecemos que nossa amostragem é pouco representativa atendendo à quantidade de utilizadores *online* presentes nas plataformas onde este foi divulgado: Facebook, Instagram, WhatsApp e via *email*. Contudo, a quantidade de respostas não tornou a pesquisa inválida, visto que a partir desta amostra foi possível recolher bastantes dados e retirar bastantes conclusões, aplicando os conhecimentos teóricos elaborados na revisão bibliográfica. Depois de analisar os *sites*

de comércio eletrónico, percebeu-se que no questionário deveriam ter sido feitas outras questões relativas a aspetos abordados nesse estudo. Outra limitação sentida foi a impossibilidade de aceder aos dados de qualquer *site* a partir da ferramenta de pesquisa da Cookiepedia, devido aos bloqueadores utilizados pelos *sites*.

Num possível futuro estudo, pretende-se fazer um estudo mais ligado aos *sites* de comércio eletrónico mais usados em Portugal, visto não ter sido possível nesta dissertação, dadas as limitações da Cookiepedia e os bloqueadores utilizados pelos *sites*.

Para estudos futuros recomenda-se que sejam abordados um maior número de inquiridos, podendo expandir a amostra e ser feita uma análise com um número maior de utilizadores online, contemplando uma população mais diversificada e utilizando um questionário mais específico em relação à recolha de dados pessoais em ambiente de comércio eletrónico, evitando assim um enviesamento das respostas. Uma vez que os utilizadores online aumentam a cada dia que passa e que este meio está constantemente em evolução e mudança, é necessária a reavaliação das conclusões agora retiradas. Da mesma forma, pretende-se também reavaliar os *sites* analisados, visto que os mesmos são atualizados consoante as necessidades dos consumidores e das atualizações do RGPD e elaborar um estudo com uma amostra de *sites* de comércio eletrónico maior.

REFERÊNCIAS

- Akter, S., & Wamba, S. F. (2016). Big data analytics in E-commerce: a systematic review and agenda for future research. *Electronic Markets*, 26(2), 173–194. <https://doi.org/10.1007/s12525-016-0219-0>
- Almeida, F. (2017). Benefits, Challenges and Tools of Big Data Management. *Journal of Systems Integration*, 8(4), 12–20. <https://doi.org/10.20470/jsi.v8i4.311>
- Apache Hadoop (2019). The Apache Software Foundation. Retrieved from <https://hadoop.apache.org/>, last accessed 2019/05/18.
- Bhadani, A. K., & Jothimani, D. (2016). Big data: Challenges, opportunities, and realities. *Effective Big Data Management and Opportunities for Implementation*, 1–24. <https://doi.org/10.4018/978-1-5225-0182-4.ch001>
- Cahn, A., Alfeld, S., Barford, P., & Muthukrishnan, S. (2016). An empirical study of web *cookies*. *25th International World Wide Web Conference, WWW 2016*, 891–901. <https://doi.org/10.1145/2872427.2882991>
- Chen, H., Chiang, R. H. L., Storey, V. C., & Robinson, J. M. (2012). *Business Intelligence Research Business Intelligence and Analytics: From Big Data To Big Impact*. 36(4), 1165–1188. www.freakonomics.com/2008/02/25/hal-varian-answers-your-questions/
- Cookiepedia (2019). How We Classify *Cookies*. Retrieved 18 March 2020, from <https://cookiepedia.co.uk/classify-cookies>
- Cookiepedia (2019). Types of *Cookies*. Retrieved 18 March 2020, from <https://cookiepedia.co.uk/types-of-cookies>
- Cookiepedia (2019). What are *cookies*?. Retrieved 17 March 2020, from <https://cookiepedia.co.uk/all-about-cookies>
- Dalfovo, M. S., Lana, R. A., & Silveira, A. (2008). Métodos Quantitativos e Qualitativos: um Resgate Teórico. *Revista Interdisciplinar Científica Aplicada*, 2(4), 1–13. http://www.unisc.br/portal/upload/com_arquivo/metodos_quantitativos_e_qualitativos_um_resgate_teorico.pdf
- Davenport, T. H., Barth, P., & Bean, R. (2012). How ‘Big Data’ Is Different. *MIT Sloan Management Review*, 54(1), 22–24.
- Demchenko, Y., De Laat, C. and Membrey, P. (2014). ‘Defining architecture components of the Big Data Ecosystem’. *Proceedings of 2014 International Conference on Collaboration Technologies and Systems (CTS)*, Minneapolis, USA, pp. 104-112.
- Diebold, F. X. (2013). A Personal Perspective on the Origin(s) and Development of “Big Data”: The Phenomenon, the Term, and the Discipline, Second Version. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.2202843>
- Diehl, Astor Antonio. (2004). *Pesquisa em ciências sociais aplicadas: métodos e técnicas*. São Paulo: Prentice Hall.
- European Commission (2017). Article 29 Data Protection Working Party: Guidelines on the right to data portability. 1–20.

- European Commission. (2016). “Factsheets on the “Right to Be Forgotten ruling (C-131/12)”.
- Félix, B. M., Tavares, E., & Cavalcante, N. W. F. (2018). Fatores críticos de sucesso para adoção de Big Data no varejo virtual: Estudo de caso do Magazine Luiza. *Revista Brasileira de Gestão de Negócios*, 20(1), 112–126. <https://doi.org/10.7819/rbgn.v20i1.3627>
- Gil, A. C. (1999). Métodos e Técnicas de Pesquisa Social. In *Métodos e técnicas de pesquisa social* (Vol. 264).
- Goldfeder, S., Kalodner, H., Reisman, D., & Narayanan, A. (2018). When the cookie meets the blockchain: Privacy risks of web payments via cryptocurrencies. *Proceedings on Privacy Enhancing Technologies*, 2018(4), 179–199. <https://doi.org/10.1515/popets-2018-0038>
- Hill, M. M. & Hill, A. (2009) Investigação por Questionário. Lisboa: Edições Sílabo.
- Ishwarappa, & Anuradha, J. (2015). A brief introduction on big data 5Vs characteristics and hadoop technology. *Procedia Computer Science*, 48(C), 319–324. <https://doi.org/10.1016/j.procs.2015.04.188>
- Jin, X., Wah, B. W., Cheng, X., & Wang, Y. (2015). Significance and Challenges of Big Data Research. *Big Data Research*, 2(2), 59–64. <https://doi.org/10.1016/j.bdr.2015.01.006>
- Katal, A., Wazid, M., & Goudar, R. H. (2013). Big data: Issues, challenges, tools and Good practices. 2013 6th International Conference on Contemporary Computing, IC3 2013, 404–409. <https://doi.org/10.1109/IC3.2013.6612229>
- Kaur, P., & Monga, A. A. (2016). Managing Big Data: A Step towards Huge Data Security. *International Journal of Wireless and Microwave Technologies*, 6(2), 10–20. <https://doi.org/10.5815/ijwmt.2016.02.02>
- Kitchin, R., & McArdle, G. (2016). What makes Big Data, Big Data? Exploring the ontological characteristics of 26 datasets. *Big Data and Society*, 3(1), 1–10. <https://doi.org/10.1177/2053951716631130>
- Kuner, C., Cate, F. H., Millard, C., & Svantesson, D. J. B. (2012). The challenge of “big data” for data protection. *International Data Privacy Law*, 2(2), 47–49. <https://doi.org/10.1093/idpl/ips003>
- Laney, D. (2001). 3D Data Management: Controlling Data Volume, Velocity, and Variety (). META Group .
- Laudon, K. C., & Traver, C. G. (2019). Comércio eletrônico 2019: business, technology, society (Thirteenth). Pearson.
- Lei n.º 46/2012 de 29 de agosto de 2012. Diário da República n.º 167, I Série. Lisboa: Assembleia da República.
- Lenz, R. (2019). Big Data: Ethics and Law. *SSRN Electronic Journal*, (September). <https://doi.org/10.2139/ssrn.3459004>
- Machado, F. N. R. (2018). Big Data: o Futuro dos Dados e Aplicações (1a ed.). São Paulo: Érica.

- Martins, G. (2008). Estudo de caso: uma reflexão sobre a aplicabilidade em pesquisa no Brasil. *Revista De Contabilidade E Organizações*, 2(2), 9-18. <https://doi.org/10.11606/rco.v2i2.34702>
- McAfee, A., & Brynjolfsson, E. (2012). Big Data: The Management revolution. *Harvard Business Review*, (October 2012), 59-68. Retrieved from papers3://publication/uuid/9D9604FA-4931-4805-ABAA-A725E4221B57
- Moniruzzaman, A. B. M., & Hossain, S. A. (2013). *NoSQL Database: New Era of Databases for Big data Analytics - Classification, Characteristics and Comparison*. May 2014. <http://arxiv.org/abs/1307.0191>
- Moysés, G. L. R., & Moori, R. G. (2007). Coleta de dados para a pesquisa acadêmica: um estudo sobre a elaboração, a validação e a aplicação eletrônica de questionário. XXVII Encontro Nacional de Engenharia de Produção (pp. 1-10). Foz do Iguaçu-PR: XXVII Encontro Nacional de Engenharia da Produção. Recuperado em: http://abepro.org.br/biblioteca/ENEGEP2007_TR660483_9457.pdf
- Cookiepedia. (2019). *Online Tracking Technologies*. Retrieved 19 March 2020, from <https://cookiepedia.co.uk/Content/OtherTrackingTechnologies>
- Parlamento Europeu, Conselho Europeu. Regulamento (UE) 2016/679 do Parlamento Europeu e do Conselho, de 27 de abril de 2016, relativo à proteção das pessoas singulares no que diz respeito ao tratamento de dados pessoais e à livre circulação desses dados e que revoga a Diretiva 95/46/CE (Regulamento Geral sobre a Proteção de Dados). JOUE. 2016;L(119):1-88.
- Pierson, J., & Heyman, R. (2011). Social media and *cookies*: Challenges for online privacy. *Info*, 13(6), 30–42. <https://doi.org/10.1108/14636691111174243>
- Porres, E. L. (2013). The Big Potential of Big Data. *Forbes Insights* .
- Prota, L., de Abreu, ana C. S. B., Suguihiro, V. L., Machado, E. M., Solci, S. M., Rizotti, M. L. A., Rodrigues, E. A. P., & Souza, C. G. (1999). Serviço Social em revista. 2, 2, 135–149.
- Raheem, N. (2019). Big Data: A Tutorial-Based Approach. (T. & F. Group, Ed.) (1 ed.). <https://doi.org/10.1201/9780429060939> Review, 54(1), 43-46.
- Sharda, R., Delen, D., & Turban, E. (2017). *Business Intelligence, Analytics and Data Science: A Managerial Perspective* (4ed ed.). Prentice Hall Press Upper Saddle River, NJ, US.
- Sivarajah, U., Kamal, M. M., Irani, Z., & Weerakkody, V. (2017). Critical analysis of Big Data challenges and analytical methods. *Journal of Business Research*, 70, 263–286. <https://doi.org/10.1016/j.jbusres.2016.08.001>
- Sousa, M. J., & Baptista, C. S. (2011). *Como fazer investigação, dissertações, teses e relatórios*. Lisboa: Pactor.
- Taurion, Cezar. Big data. Rio de Janeiro: Brasport, 2013. ePUB.
- Yin, R. K. (2001). *Estudo de caso: planejamento e métodos*. Porto Alegre: Bookman.

- Yin, R. K. (2018). Case study research and applications: Design and methods. In *Journal of Hospitality & Tourism Research* (6 ed., Vol. 53). <https://doi.org/10.1177/109634809702100108>
- Zhang, F., Liu, M., Gui, F., Shen, W., Shami, A., & Ma, Y. (2015). A distributed frequent itemset mining algorithm using Spark for Big Data analytics. *Cluster Computing*, 18(4), 1493–1501.

Apêndice I – Questionário



INSTITUTO
SUPERIOR
DE CONTABILIDADE
E ADMINISTRAÇÃO
DO PORTO

QUESTIONÁRIO

Esclarecimentos sobre a pesquisa

Este questionário foi elaborado como suporte para a realização de uma dissertação do Mestrado em Negócio Eletrónico no Instituto Superior de Contabilidade e Administração do Porto sob orientação da Doutora Ana Azevedo. Esta dissertação pretende perceber a utilização dos Big Data no contexto do comércio eletrónico, quer do ponto de vista dos utilizadores, quer do ponto de vista das organizações.

O objetivo deste questionário visa comprovar se os utilizadores *online* têm consciência que os seus dados pessoais são recolhidos pelos *sites* de comércio eletrónico, percebendo quais os motivos para a sua partilha, qual o seu comportamento perante as técnicas de recolha de dados e o nível de conhecimento perante o Regulamento Geral de Proteção de Dados.

Os dados desta pesquisa são confidenciais e sua identidade não será revelada. Desde já agradeço a sua colaboração e saliento que a sua participação é fundamental para o sucesso deste estudo.

Este questionário demora cerca de 5 minutos a ser respondido.

CARACTERIZAÇÃO DA AMOSTRA

1. Indique a sua faixa etária.

2. Indique o seu sexo.

☐ Masculino ☐ Feminino

3. Indique o seu grau de escolaridade mais elevado.

☐ Ensino básico – 1º Ciclo

☐ Ensino básico – 2º Ciclo

☐ Ensino secundário

☐ Licenciatura

☐ Mestrado

☐ Doutoramento

☐ Outro _____

4. Que tipo de dispositivo mais utiliza para aceder à *Internet*?

☐ Computador

- ☐ Tablet
- ☐ Smartphone
- ☐ Outro. Qual? _____

NÍVEL DE CONHECIMENTO DO UTILIZADOR <i>ONLINE</i> EM RELAÇÃO AO CONCEITO DE DADOS PESSOAIS

5. Das opções abaixo, indique aquelas que considera serem dados pessoais.

Nota: Pode assinalar tantas opções quantas as que considerar corretas

- ☐ Nome
- ☐ Morada
- ☐ Número de telemóvel
- ☐ Email
- ☐ Dados de contas bancárias
- ☐ Género
- ☐ Idade
- ☐ Ocupação
- ☐ Localização
- ☐ Gostos (*likes*)
- ☐ Impressões digitais
- ☐ Nível de Educação
- ☐ Transações financeiras
- ☐ Tipo de navegador

6. Costuma fornecer os seus dados pessoais em contexto de comércio eletrónico?

- ☐ Sim ☐ Não

7. Caso tenha respondido SIM, indique as 3 principais razões para fornecer que o levam a fornecer os seus dados pessoais em contexto de comércio eletrónico.

- ☐ Fazer um pagamento
- ☐ Obter um serviço
- ☐ Poupar tempo na próxima visita
- ☐ Obter um serviço gratuito
- ☐ Conseguir a entrega de uma compra
- ☐ Nunca forneço dados *online*
- ☐ Outras. Quais? _____

AVALIAÇÃO DO COMPORTAMENTO DO UTILIZADOR *ONLINE* PERANTE A PARTILHA DOS SEUS DADOS PESSOAIS

8. Que controlo sente que tem sobre a informação que disponibiliza *online*?

- ☐ Total
- ☐ Parcial
- ☐ Nenhum
- ☐ Depende do *site* ou aplicação
- ☐ Não sei

9. Assinale a resposta com que mais se identifica:

	SIM	NÃO
Evita alguns <i>sites</i> de comércio eletrónico por receio de que a atividade seja monitorizada?		
Mudou as definições de privacidade do navegador que utiliza?		
Utiliza software que evita que as suas atividades <i>online</i> sejam monitorizadas?		
Quando acede a um <i>site</i> , tem por hábito aceitar os <i>cookies</i> ?		

NÍVEL DE CONHECIMENTO PERANTE OS SEUS DIREITOS INERENTES NO RGDP

10. Tem conhecimento do que é o Regulamento Geral de Proteção de Dados?

- ☐ Sim ☐ Não

11. Caso tenha respondido SIM, relativamente ao Regulamento Geral de Proteção de Dados, indique a opção que considera correta:

	SIM	NÃO
O Regulamento aplica-se em <i>sites</i> de comércio eletrónico fora da União Europeia.		
O RGPD obriga as empresas a registar os dados pessoais dos cidadãos num formato que permita que sejam enviados para outra empresa, facilitando a sua migração. (Direito à Portabilidade dos Dados)		
Os cidadãos podem exigir que os seus dados pessoais sejam eliminados. (Direito a ser Esquecido)		

NÍVEL DE CONCORDÂNCIA PERANTE AS OPÇÕES QUE OS SITES OFERECEM PARA A SUA UTILIZAÇÃO

12. Considera aceitável:

	SIM	NÃO
Que a sua atividade <i>online</i> seja monitorizada, em troca de acesso livre a um <i>site</i> ?		
Que os <i>sites</i> de comércio eletrónico utilizem os seus dados como utilizador para benefício próprio?		
Que os <i>sites</i> de comércio eletrónico partilhem os seus dados como utilizador sem a sua permissão?		
Pagar para não ser monitorizado enquanto navega num <i>site</i> ?		

13. Numa escala de 1 a 5, sendo que 1 representa discordo totalmente e 5 representa concordo totalmente, avalie as seguintes afirmações:

	1	2	3	4	5
Fornecer informação pessoal <i>online</i> faz cada vez mais parte da vida moderna.					
Sinto que tenho de fornecer cada vez mais informação pessoal.					
Sinto-me invadido ao fornecer informação pessoal.					
Importo-me de fornecer informação pessoal em troca de serviços <i>online</i> gratuitos, como o email.					
Não há alternativa para obter produtos ou serviços sem que tenha de fornecer dados pessoais.					

14. Pretende deixar algum comentário adicional?

Anexo I – Cookies utilizados pelo continente.pt

The screenshot shows the Continente.pt website with a modal window titled "Política de Cookies" (Cookie Policy) displayed in the center. The modal contains a table with two columns: "Nome do Cookie" (Cookie Name) and "Finalidade" (Purpose). The table lists three cookies: ASP.NET_SessionId, __CommerceAnonymousShopper_ef77e72d-62b9-4b0f-8113-d111c9d6d7ce_Internet, and AnonymousBasket. Below the table is a "Fechar" (Close) button. The background of the website is dimmed, showing the top navigation bar with links like "CONTINENTE", "Vida + Saudável", "CASA JARDIM", "brinquedos", "animais Vet ZU", "Livraria & Papelaria", "well's", "CONTINENTE Negócios", and "ENTREGAZERO". The footer of the website includes links like "Robalo", "Marca da Semana", "Promoções", "Dia Mundial da Criança", "Covid 19", "Ar Livre", "PORTUGUÊS", "MAPA DO SITE", "LOJAS PERTO DE SI", "CONTACTOS", "MODELO CONTINENTE HIPERMERCADOS, S.A.", "POLÍTICA DE SERVIÇOS", and "EMPREGO".

Nome do Cookie	Finalidade
ASP.NET_SessionId	Preserva o estado da sessão do visitante entre pedidos de páginas.
__CommerceAnonymousShopper_ef77e72d-62b9-4b0f-8113-d111c9d6d7ce_Internet	Identifica o utilizador no Continente Online Commerce System.
AnonymousBasket	Mantém o registo dos produtos adicionados ao carrinho por um utilizador anónimo.

Fechar

O nosso website utiliza cookies para melhorar e personalizar a sua experiência de navegação. Ao continuar a navegar está a consentir a utilização de cookies. [Saiba mais.](#)

Email

Palavra-passe

Entrar

Novo Registo | Perdeu a Palavra-passe?

CONTINENTE

Vida + Saudável

CASA JARDIM

brinquedos

animais Vet ZU

LIVRARIA & Papelaria

well's

CONTINENTE Negócios

ENTREGAZERO

O que procura?

NewsletterFolhetosAjudaCartão ContinenteCartão Dá

REDES SOCIAIS Sig...

FacebookTwitterInstagramLinkedInWhatsApp

Política de Cookies

CampaignHistory	Informação persistente sobre produtos do Comm System.
CheckForConsentsVersionsChangesAlreadyDoneKey	Mantém o registo da apresentação da mensagem/consentimento do RGPD após mudanç versão.
cPrompt_useCookies	Permite-nos saber se o utilizador aceitou cookies site.
cpup	Mantém o registo da exibição da mensagem dos cupões.

Fechar

Robalo : Marca da Semana : Promoções : Dia Mundial da Criança : Covid 19 : Ar Livre

PORTUGUÊS ▾ ▾

MAPA DO SITELOJAS PERTO DE SICONTACTOSMODELO CONTINENTE HIPERMERCADOS, S.A.POLÍTICA DE SERVIÇOSEMPREGO

PARTILHA

© COPYRIGHT 2020 MODELO CONTINENTE HIPERMERCADOS, S.A. TODOS OS DIREITOS RESERVADOS

O nosso website utiliza cookies para melhorar e personalizar a sua experiência de navegação. Ao continuar a navegar está a consentir a utilização de cookies. Saiba mais.


Email


Palavra-passe


Entrar


Novo Registo | Perdeu a Palavra-passe?


 CONTINENTE

 Vida + Saudável

 CASA JARDIM


 brinquedos'

 animais Vet ZU

 LIVRARIA & Papelaria

 well's

 CONTINENTE Negócios

 ENTREGAZERO

O que procura?

Newsletter

Folhetos

Ajuda

Cartão Continente

Cartão Dá

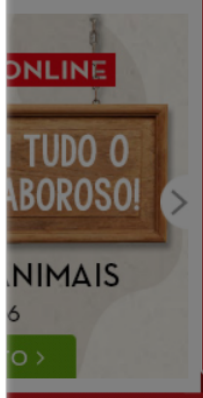
REDES SOCIAIS

Siga-nos

para a sua vida. Ver apps







Robalo

Marca da Semana

Promoções

Dia Mundial da Criança

Covid 19

Ar Livre

PORTUGUÊS

MAPA DO SITE

LOJAS PERTO DE SI

CONTACTOS

MODELO CONTINENTE HIPERMERCADOS, S.A.

POLÍTICA DE SERVIÇOS

EMPREGO

Partilhar

© COPYRIGHT 2020 MODELO CONTINENTE HIPERMERCADOS, S.A. TODOS OS DIREITOS RESERVADOS

O nosso website utiliza cookies para melhorar e personalizar a sua experiência de navegação. Ao continuar a navegar está a consentir a utilização de cookies. [Saiba mais.](#)

Política de Cookies

f5avrbbbbbbbbbbbbbb	Usado pelo balanceador de carga para dirigir o tráfego para diferentes servidores.
ModalLayerCheck_images	Mantém o registo da exibição da mensagem de aniversário do Continente Online.
MSCSPProfile	Recipiente e rastreador do perfil de informação do Microsoft Commerce Server.
__cfduid	Usada pela rede de conteúdos Cloudflare para registar o tráfego de confiança da web

Fechar